

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2002, 13:03:48 ; Search time 105.16 Seconds
(without alignments)
20.083 Million cell updates/sec

Title: US-09-627-383-1
Perfect score: 29
Sequence: 1 LEPRAS 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues
Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US10_COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US10_COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length DB	ID	Description
1	29	100.0	6	20	US-09-627-383-1
2	29	100.0	67	20	US-60-146-315-866
3	29	100.0	75	1	PCT-US01-14827-12024
4	29	100.0	108	1	PCT-US01-14827-14753
5	29	100.0	119	1	PCT-US01-08656-7642
6	29	100.0	153	21	US-09-758-472-5581
7	29	100.0	256	12	US-08-878-322-11

8	29	100.0	256	14	US-09-099-295-11	Sequence 11, Appl
9	29	100.0	256	22	US-09-852-472-11	Sequence 11, Appl
10	29	100.0	258	21	US-09-758-440-688	Sequence 688, App
11	29	100.0	421	23	US-09-902-540-13318	Sequence 13318, A
12	29	100.0	439	16	US-09-270-767-57034	Sequence 57034, A
13	29	100.0	439	16	US-09-270-849B-180914	Sequence 180914, A
14	29	100.0	471	26	US-60-206-047-406	Sequence 406, App
15	29	100.0	634	26	US-60-171-625-531	Sequence 531, App
16	29	100.0	634	26	US-60-173-464-26180	Sequence 26180, A
17	29	100.0	635	20	US-09-614-150-34179	Sequence 34179, A
18	29	100.0	635	26	US-60-191-637-33763	Sequence 33763, A
19	29	100.0	635	26	US-60-191-681-26617	Sequence 26617, A
20	29	100.0	635	26	US-60-219-005-312	Sequence 312, App
21	29	100.0	637	1	PCT-US01-08631-52925	Sequence 52925, A
22	29	100.0	666	16	US-09-270-767-41791	Sequence 41791, A
23	29	100.0	682	1	PCT-US00-25340-1	Sequence 1, Appl
24	29	100.0	708	12	US-08-878-322-8	Sequence 8, Appl
25	29	100.0	708	12	US-09-099-295-8	Sequence 8, Appl
26	29	100.0	708	22	US-09-852-472-8	Sequence 8, Appl
27	29	100.0	966	18	US-09-488-725A-1824	Sequence 1824, App
28	29	100.0	1013	18	US-09-488-725A-1823	Sequence 1823, App
29	29	100.0	1021	26	US-60-167-217-7726	Sequence 7726, App
30	29	100.0	1025	18	US-09-488-725A-1825	Sequence 1825, App
31	29	100.0	1060	26	US-60-173-464-6181	Sequence 6181, App
32	29	100.0	1077	20	US-09-614-150-10764	Sequence 10764, A
33	29	100.0	1077	26	US-60-167-217-10786	Sequence 10786, A
34	29	100.0	1077	26	US-60-173-464-8743	Sequence 8743, App
35	29	100.0	1077	26	US-60-191-637-10796	Sequence 10796, App
36	29	100.0	1077	26	US-60-191-681-8465	Sequence 8465, App
37	29	100.0	1219	1	PCT-US99-14482-4	Sequence 4, Appl
38	29	100.0	1219	17	US-09-344-624-4	Sequence 4, Appl
39	29	100.0	1253	1	PCT-US99-19663-2	Sequence 2, Appl
40	29	100.0	1253	17	US-09-388-089-2	Sequence 2, Appl
41	29	100.0	2359	17	US-09-346-794-26	Sequence 26, Appl
42	29	100.0	2359	20	US-09-611-257-26	Sequence 26, Appl
43	29	100.0	2548	22	US-09-851-682-1	Sequence 1, Appl
44	29	100.0	2548	22	US-09-851-682A-1	Sequence 1, Appl
45	29	100.0	2887	21	US-09-742-153-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-627-383-1
; Sequence 1, Application US/09627383.
; GENERAL INFORMATION:
; APPLICANT: Paul T. Matsudaira
; APPLICANT: Daniel J. Enrllich
; APPLICANT: Oihui Zhong
; APPLICANT: Yelena Freyzen
; TITLE OF INVENTION: Affinity Fluorescent Proteins and Uses
; FILE REFERENCE: 0399.1211-001
; CURRENT APPLICATION NUMBER: US/09/627,383
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/146,438
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/061,801
; PRIOR FILING DATE: 1997-10-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hexapeptide
US-09-627-383-1
Query Match 100.0%; Score 29; DB 20; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.2e+06;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
|||||
Db 1 LEPRAS 6

RESULT 2
US-60-146-315-866
; Sequence 866, Application US/60146315
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED KINASE PROTEINS, NUCLEIC ACID
; FILE REFERENCE: C1000064
; CURRENT APPLICATION NUMBER: US/60/146,315
; CURRENT FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 1086
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 866
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Drosophila
US-60-146-315-866

Query Match 100.0%; Score 29; DB 26; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
|||||
Db 16 LEPRAS 21

RESULT 3
PCT-US01-14827-12024
; Sequence 12024, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 12024
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (30)..(41)
; OTHER INFORMATION: PROTEIN SPORULATION REPEAT PRECU domain identified by
; OTHER INFORMATION: eMATRIX, accession number P02283B, P-value=7.857e-09, raw score
; PCT-US01-14827-12024

Query Match 100.0%; Score 29; DB 1; Length 75;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
|||||
Db 63 LEPRAS 68

RESULT 4
PCT-US01-14827-14753
; Sequence 14753, Application PC/TUS0114827

; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 14753
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(108)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-14827-14753

Query Match 100.0%; Score 29; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
|||||
Db 35 LEPRAS 40

RESULT 5
PCT-US01-08656-7642
; Sequence 7642, Application PC/TUS0108656
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: PCT/US01/08656
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 7642
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-08656-7642

Query Match 100.0%; Score 29; DB 1; Length 119;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
|||||
Db 81 LEPRAS 86

RESULT 6
US-09-758-472-5581
; Sequence 5581, Application US/09758472
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PH001
; CURRENT APPLICATION NUMBER: US/09/758,472
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628

PRIOR FILING DATE: 2000-02-04
NUMBER OF SEQ ID NOS: 9632
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5581
LENGTH: 153
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (115)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (122)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (123)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-472-5581

Query Match 100.0%; Score 29; DB 21; Length 153;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
Db 54 LEPRAS 59

RESULT 7
US-08-878-322-11
Sequence 11, Application US/08878322
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Jelinek, Laura J.
APPLICANT: Whitmore, Theodore E.
APPLICANT: Blumberg, Hal
TITLE OF INVENTION: MAMMALIAN ZNEU1
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zymogenetics, Inc.
STREET: 1201 Eastlake Ave. East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,322
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Paul G
REGISTRATION NUMBER: 32,743
REFERENCE/DOCKET NUMBER: 97-28X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6627
TELEFAX: 206-442-6678
TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
ORIGINAL SOURCE:

ORGANISM: HSMHC3W5A-6 LIKE DOMAIN
US-08-878-322-11

Query Match 100.0%; Score 29; DB 12; Length 256;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
Db 146 LEPRAS 151

RESULT 8
US-09-099-295-11
Sequence 11, Application US/09099295
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Jelinek, Laura J.
APPLICANT: Whitmore, Theodore E.
APPLICANT: Blumberg, Hal
TITLE OF INVENTION: MAMMALIAN NEURO-GROWTH FACTOR LIKE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zymogenetics
STREET: 1201 Eastlake Ave East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/099,295
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Paul G
REGISTRATION NUMBER: 32,743
REFERENCE/DOCKET NUMBER: 97-28
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6627
TELEFAX: 206-442-6678
TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-099-295-11

Query Match 100.0%; Score 29; DB 14; Length 256;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
Db 146 LEPRAS 151

RESULT 9
US-09-852-472-11

```
; Sequence 11, Application US/09852472
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Jelinek, Laura J.
; TITLE OF INVENTION: Mammalian Neuro-Growth Factor Like
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 97-28C1
; CURRENT APPLICATION NUMBER: US/09/852,472
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 09/039,295
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/050,143
; PRIOR FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-852-472-11
```

```
Query Match          100.0%; Score 29; DB 22; Length 256;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 LEPRAS 6
    |||||
Db 146 LEPRAS 151
```

```
RESULT 10
US-09-758-440-688
; Sequence 688, Application US/09758440
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM015
; CURRENT APPLICATION NUMBER: US/09/758,440
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 832
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 688
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-440-688
```

```
Query Match          100.0%; Score 29; DB 21; Length 258;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 LEPRAS 6
    |||||
Db 67 LEPRAS 72
```

```
RESULT 11
US-09-902-540-13318
; Sequence 13318, Application US/09902540
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
```

```
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13318
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13318
```

```
Query Match          100.0%; Score 29; DB 23; Length 421;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 LEPRAS 6
    |||||
Db 406 LEPRAS 411
```

```
RESULT 12
US-09-270-767-57034
; Sequence 57034, Application US/09270767
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57034
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-57034
```

```
Query Match          100.0%; Score 29; DB 16; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 LEPRAS 6
    |||||
Db 196 LEPRAS 201
```

```
RESULT 13
US-09-270-849B-180914
; Sequence 180914, Application US/09270849B
; GENERAL INFORMATION:
; APPLICANT: Swimmer et al.
; TITLE OF INVENTION: Insect genome survey devices
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/270,849B
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 195450
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 180914
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-180914
```

Query Match 100.0%; Score 29; DB 16; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
|||||
DB 196 LEPRAS 201

RESULT 14

US-60-206-047-406
; Sequence 406, Application US/60206047
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PHASE I DRUG-METABOLIZING
; PROTEINS; NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PHASE I PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO00570
; CURRENT APPLICATION NUMBER: US/60/206,047
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 456
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 406
; LENGTH: 471
; TYPE: PRT
; ORGANISM: HUMAN
US-60-206-047-406

Query Match 100.0%; Score 29; DB 26; Length 471;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
|||||
DB 282 LEPRAS 287

RESULT 15

US-60-171-625-531
; Sequence 531, Application US/60171625
; GENERAL INFORMATION:
; APPLICANT: Skupski, Marian
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS THAT SHOW
; TITLE OF INVENTION: HIGH HOMOLOGY TO KNOWN HUMAN DISEASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00179
; CURRENT APPLICATION NUMBER: US/60/171,625
; CURRENT FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 579
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 531
; LENGTH: 634
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-171-625-531

Query Match 100.0%; Score 29; DB 26; Length 634;
Best Local Similarity 100.0%; Pred. No. 2.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
|||||
DB 26 LEPRAS 31

Search completed: June 6, 2002, 13:07:20
Job time: 212 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2002, 13:03:48 ; Search time 12.67 Seconds
(without alignments)
27.907 Million cell updates/sec

Title: US-09-627-383-1

Perfect score: 29

Sequence: 1 LEPRAS 6

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 186664 seqs, 58930901 residues

Total number of hits satisfying chosen parameters: 186664

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing First 45 summaries

Database :

Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	89.7	61	5	US-09-540-209B-9237
2	26	89.7	153	6	US-10-121-062-564
3	26	89.7	406	1	PCT-US02-11152-21
4	26	89.7	485	1	PCT-US02-11152-20
5	26	89.7	8360	6	US-10-132-134-34
6	25	86.2	24	1	PCT-US02-07282-6
7	25	86.2	24	6	US-10-114-500-6
8	25	86.2	58	5	US-09-855-604-538
9	25	86.2	177	6	US-10-113-872-1678
10	25	86.2	214	5	US-09-540-209B-5669
11	25	86.2	241	5	US-09-835-625-23288
12	25	86.2	283	5	US-09-895-913A-28
13	25	86.2	287	5	US-09-935-625-23287
14	25	86.2	314	6	US-10-113-872-1863
15	25	86.2	367	7	US-60-377-714-52
16	25	86.2	439	1	PCT-US02-07282-4
17	25	86.2	439	6	US-10-114-500-4
18	25	86.2	475	5	US-09-935-625-13383
19	25	86.2	475	5	US-09-935-625-21538
20	25	86.2	475	5	US-09-935-625-30024
21	25	86.2	535	5	US-09-935-625-13382
22	25	86.2	535	5	US-09-935-625-21537
23	25	86.2	535	5	US-09-935-625-30023
24	25	86.2	551	5	US-09-835-625-11616
25	25	86.2	551	5	US-09-935-625-19542
26	25	86.2	551	5	US-09-935-625-28022

27	25	86.2	590	5	US-09-540-209B-6352	Sequence 6352, Ap
28	25	86.2	611	5	US-09-935-625-11615	Sequence 11615, A
29	25	86.2	611	5	US-09-935-625-19541	Sequence 19541, A
30	25	86.2	611	5	US-09-935-625-28021	Sequence 28021, A
31	25	86.2	619	5	US-09-573-655B-1686	Sequence 1686, Ap
32	25	86.2	630	5	US-09-935-625-13381	Sequence 13381, A
33	25	86.2	630	5	US-09-935-625-21536	Sequence 21536, A
34	25	86.2	630	5	US-09-935-625-30022	Sequence 30022, A
35	25	86.2	660	5	US-09-935-625-17623	Sequence 17623, A
36	25	86.2	668	5	US-09-935-625-17045	Sequence 17045, A
37	25	86.2	700	1	PCT-US02-13142-1117	Sequence 1117, Ap
38	25	86.2	700	6	US-10-128-714-3117	Sequence 3117, Ap
39	25	86.2	706	5	US-09-935-625-11614	Sequence 11614, A
40	25	86.2	706	5	US-09-935-625-19540	Sequence 19540, A
41	25	86.2	706	5	US-09-935-625-28020	Sequence 28020, A
42	25	86.2	798	5	US-09-935-625-17622	Sequence 17622, A
43	25	86.2	806	5	US-09-935-625-17044	Sequence 17044, A
44	25	86.2	845	5	US-09-890-709-4	Sequence 4, Appl
45	25	86.2	877	5	US-09-573-655B-150	Sequence 150, App

ALIGNMENTS

RESULT 1
US-09-540-209B-9237
; Sequence 9237, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
; FILE REFERENCE: 2709, 1001-001
; CURRENT APPLICATION NUMBER: US/09/540, 209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 9237
; LENGTH: 61
; TYPE: PRT
; ORGANISM: B.fragilis
; US-09-540-209B-9237

Query Match 89.7%; Score 26; DB 5; Length 61;
Best Local Similarity 83.3%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
Db 1 LEPRAS 6

RESULT 2
US-10-121-062-564
; Sequence 564, Application US/10121062
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1c1
; CURRENT APPLICATION NUMBER: US/10/121,062
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612

SEQ ID NO 564
LENGTH: 153
TYPE: PRT
ORGANISM: Homo Sapien
US-10-121-062-564

Query Match 89.7%; Score 26; DB 6; Length 153;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
1111:1
DB 145 LEPRAS 150

RESULT 3
PCT-US02-11152-21
Sequence 21, Application PC/TUS0211152
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: AZIMZAI, Yalda
APPLICANT: AU-YOUNG, Janice K.
APPLICANT: BATRA, Sajeiv
APPLICANT: BAUGHN, Mariah R.
APPLICANT: BECHA, Shanya D.
APPLICANT: BOROMSKY, Mark L.
APPLICANT: BUFORD, Neil
APPLICANT: DING, Li
APPLICANT: ELLIOTT, Vicki S.
APPLICANT: EMERLING, Brooke M.
APPLICANT: GANDHI, Ameena R.
APPLICANT: GIETZEN, Kimberly J.
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: HAFALIA, April J.A.
APPLICANT: HONCHELL, Cynthia D.
APPLICANT: LAL, Preeti G.
APPLICANT: LEE, Soo Yeun
APPLICANT: LU, Dyung Aina M.
APPLICANT: ARVIZU, Chandra S.
APPLICANT: RAKKUMAR, Jayalaxmi
APPLICANT: REDDY, Roopa
APPLICANT: SANJAMMALA, Madhu, M.
APPLICANT: TANG, Y. Tom
APPLICANT: WALIA, Narinder K.
APPLICANT: WANG, Yu-mei, E.
APPLICANT: WARREN, Bridget A.
APPLICANT: XU, Yuming
APPLICANT: YANG, Junming
APPLICANT: YAO, Monique G.
APPLICANT: YUE, Henry
APPLICANT: ZEBARADIAN, Yeganeh
TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
FILE REFERENCE: PI-0417 PCT
CURRENT APPLICATION NUMBER: PCT/US02/11152
CURRENT FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 60/283,110; 60/283,294; 60/286,820; 60/287,228;
60/291,662; 60/291,846; 60/293,727; 60/295,340;
60/295,263; 60/349,705
PRIOR FILING DATE: 2001-04-06; 2001-04-11; 2001-04-26; 2001-04-27;
2001-05-16; 2001-05-18; 2001-05-25; 2001-06-01;
2001-06-01; 2002-01-15
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PERL Program
SEQ ID NO 21
LENGTH: 406
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 7506195CD1
PCT-US02-11152-21

Query Match 89.7%; Score 26; DB 1; Length 406;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
11111
DB 62 LEPRAS 67

RESULT 4
PCT-US02-11152-20
Sequence 20, Application PC/TUS0211152
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: AZIMZAI, Yalda
APPLICANT: AU-YOUNG, Janice K.
APPLICANT: BATRA, Sajeiv
APPLICANT: BAUGHN, Mariah R.
APPLICANT: BECHA, Shanya D.
APPLICANT: BOROMSKY, Mark L.
APPLICANT: BUFORD, Neil
APPLICANT: DING, Li
APPLICANT: ELLIOTT, Vicki S.
APPLICANT: EMERLING, Brooke M.
APPLICANT: GANDHI, Ameena R.
APPLICANT: GIETZEN, Kimberly J.
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: HAFALIA, April J.A.
APPLICANT: HONCHELL, Cynthia D.
APPLICANT: LAL, Preeti G.
APPLICANT: LEE, Soo Yeun
APPLICANT: LU, Dyung Aina M.
APPLICANT: ARVIZU, Chandra S.
APPLICANT: RAKKUMAR, Jayalaxmi
APPLICANT: REDDY, Roopa
APPLICANT: SANJAMMALA, Madhu, M.
APPLICANT: TANG, Y. Tom
APPLICANT: WALIA, Narinder K.
APPLICANT: WANG, Yu-mei, E.
APPLICANT: WARREN, Bridget A.
APPLICANT: XU, Yuming
APPLICANT: YANG, Junming
APPLICANT: YAO, Monique G.
APPLICANT: YUE, Henry
APPLICANT: ZEBARADIAN, Yeganeh
TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
FILE REFERENCE: PI-0417 PCT
CURRENT APPLICATION NUMBER: PCT/US02/11152
CURRENT FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 60/282,110; 60/283,294; 60/286,820; 60/287,228;
60/291,662; 60/291,846; 60/293,727; 60/295,340;
60/295,263; 60/349,705
PRIOR FILING DATE: 2001-04-06; 2001-04-11; 2001-04-26; 2001-04-27;
2001-05-16; 2001-05-18; 2001-05-25; 2001-06-01;
2001-06-01; 2002-01-15
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PERL Program
SEQ ID NO 20
LENGTH: 485
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 4558650CD1
PCT-US02-11152-20

Query Match 89.7%; Score 26; DB 1; Length 485;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6

Db 62 LQPRAS 67
|:|||||

RESULT 5
US-10-132-134-34
; Sequence 34, Application US/10132134
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Tang, Xianshu
; APPLICANT: Staifa, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES
; FILE REFERENCE: 3012-205
; CURRENT APPLICATION NUMBER: US/10/132,134
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 8360
; TYPE: PRT
; ORGANISM: Streptomyces amphibiosporus
US-10-132-134-34

Query Match 89.7%; Score 26; DB 6; Length 8360;
Best Local Similarity 83.3%; Pred. No. 5.8e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPRAS 6
|:|||||
Db 6398 LEPRAS 6403

RESULT 6
PCT-US02-07282-6
; Sequence 6, Application PC/TUS0207282
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; APPLICANT: Godbole, Shubhada D
; APPLICANT: Boyle, Bryan J
; APPLICANT: Arterburn, Matthew C
; APPLICANT: Liu, Chenghua
; APPLICANT: Tang, Y Tom
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO FIBULIN-LIKE POLYPEPTIDES AND
; FILE REFERENCE: 24059-040-061
; CURRENT APPLICATION NUMBER: PCT/US02/07282
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/802,704
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-07282-6

Query Match 86.2%; Score 25; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EPRAS 6
|:|||||
Db 20 EPRAS 24

RESULT 7
US-10-114-500-6
; Sequence 6, Application US/10114500
; GENERAL INFORMATION:
; APPLICANT: Godbole, Shubhada
; APPLICANT: Boyle, Bryan J
; APPLICANT: Arterburn, Matthew C
; APPLICANT: Liu, Chenghua
; APPLICANT: Tang, Y Tom
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO FIBULIN-LIKE POLYPEPTIDES AN
; FILE REFERENCE: HYS-400CN
; CURRENT APPLICATION NUMBER: US/10/114,500
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: US 09/802,704
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-500-6

Query Match 86.2%; Score 25; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EPRAS 6
|:|||||
Db 20 EPRAS 24

RESULT 8
US-09-855-604-538
; Sequence 538, Application US/09855604
; GENERAL INFORMATION:
; APPLICANT: GICQUEL, BRIGITTE
; APPLICANT: PORTNOI, DENIS
; APPLICANT: LIM, ENG-MONG
; APPLICANT: PELICIC, VLADIMIR
; APPLICANT: GUIGUENO, AGNES
; APPLICANT: GOGUET DE LA SALMONIERE, YVES
; TITLE OF INVENTION: POLYPEPTIDE NUCLEIC SEQUENCES EXPORTED FROM MYCOBACTERIA,
; TITLE OF INVENTION: VECTORS COMPRISING SAME AND USES FOR DIAGNOSING AND
; FILE REFERENCE: 03715.0062-01000
; CURRENT APPLICATION NUMBER: US/09/855,604
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/485,536
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: PCT/FR98/01813
; PRIOR FILING DATE: 1998-08-14
; PRIOR APPLICATION NUMBER: FR 97 10404
; PRIOR FILING DATE: 1997-08-14
; PRIOR APPLICATION NUMBER: FR 97 11325
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 935
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 538
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-855-604-538

```
Query Match      86.2%; Score 25; DB 5; Length 58;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EPRAS 6
    |||||
Db 4 EPRAS 8

RESULT 9
US-10-113-872-1678
; Sequence 1678, Application US/10113872
; GENERAL INFORMATION:
; APPLICANT: Watanabe, Yoshinhiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C19
; CURRENT APPLICATION NUMBER: US/10/113,872
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1678
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-113-872-1678

Query Match      86.2%; Score 25; DB 6; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EPRAS 6
    |||||
Db 20 EPRAS 24

RESULT 10
US-09-540-209B-5669
; Sequence 5669, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 5669
; LENGTH: 214
; TYPE: PRT
; ORGANISM: B.fragilis
US-09-540-209B-5669

Query Match      86.2%; Score 25; DB 5; Length 214;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPPA 5
    |||||
Db 97 LEPPA 101

RESULT 11
US-09-935-625-23288
; Sequence 23288, Application US/09935625

; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; TITLE OF INVENTION: MODULATING VARIOUS RESPONSES
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 23288
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..241
; OTHER INFORMATION: Ceres Seq. ID no. 2028609
US-09-935-625-23288

Query Match      86.2%; Score 25; DB 5; Length 241;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPPA 5
    |||||
Db 77 LEPPA 81

RESULT 12
US-09-895-913A-28
; Sequence 28, Application US/09895913A
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Comen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding Novel Helicobacter Polypeptides in the Helicobacter
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-895-913A-28

Query Match      86.2%; Score 25; DB 5; Length 283;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPPA 5
    |||||
Db 140 LEPPA 144

RESULT 13
US-09-935-625-23287
; Sequence 23287, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
```

```

; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 23287
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..287
; OTHER INFORMATION: Ceres Seq. ID no. 2028608
US-09-935-625-23287

```

```

Query Match      86.2%; Score 25; DB 5; Length 287;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 LEPPA 5
      |||||
Db      123 LEPPA 127

```

```

RESULT 14
US-10-113-872-1863
; Sequence 1863, Application US/10113872
; GENERAL INFORMATION:
; APPLICANT: Matanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C19
; CURRENT APPLICATION NUMBER: US/10/113,872
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1863
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-113-872-1863

```

```

Query Match      86.2%; Score 25; DB 6; Length 314;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 EPPAS 6
      |||||
Db      157 EPPAS 161

```

```

RESULT 15
US-60-377-714-52
; Sequence 52, Application US/60377714
; GENERAL INFORMATION:
; APPLICANT: Ward, Teresa R
; APPLICANT: Mao, Mao
; APPLICANT: Linsley, Peter S
; TITLE OF INVENTION: LXR Ligand Induced Genes and Proteins
; FILE REFERENCE: R02-016-200
; CURRENT APPLICATION NUMBER: US/60/377,714
; CURRENT FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-377-714-52

```

```

Query Match      86.2%; Score 25; DB 7; Length 367;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 LEPPA 5
      |||||
Db      132 LEPPA 136

```

```

Search completed: June 6, 2002, 13:04:51
Job time: 63 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2002, 13:03:48 ; Search time 30 Seconds
(without alignments)
22.215 Million cell updates/sec

Title: US-09-627-383-1
Perfect score: 29
Sequence: 1 LEPRAS 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq_032802:*

- 1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	6	22	AA870246
2	29	100.0	119	22	AAU31081
3	29	100.0	125	22	AAU39701
4	29	100.0	256	20	AAW88386
5	29	100.0	526	22	AAW88386
6	29	100.0	635	22	AAW88386
7	29	100.0	637	22	AAW88386
8	29	100.0	708	22	AAW88386
9	29	100.0	966	22	AAW88386
10	29	100.0	1013	22	AAW88386
11	29	100.0	1025	22	AAW88386

12	29	100.0	1077	22	ABE61324	Drosophila melanog
13	29	100.0	1219	21	AAV78795	Human anti-tu-2 (
14	29	100.0	2359	21	AAE66476	Rat alpha-TH calci
15	29	100.0	2548	20	AAV05781	Human myosin IXa.
16	29	100.0	3096	20	AAE67770	FANCIP3 protein fr
17	27	93.1	147	20	AAV27779	Human secreted pro
18	26	89.7	15	21	AAV70693	E. coli antibody d
19	26	89.7	18	20	AAV05366	HBV specific singl
20	26	89.7	30	22	ABR11870	Human secreted pro
21	26	89.7	30	22	AAO06368	Human polypeptide
22	26	89.7	31	22	AAO06368	Human polypeptide
23	26	89.7	32	22	AAO27436	Human polypeptide
24	26	89.7	32	22	AAO10760	Human polypeptide
25	26	89.7	33	22	AAO08563	Human polypeptide
26	26	89.7	34	22	AAW90091	Human immune/haema
27	26	89.7	35	22	AAO06223	Human polypeptide
28	26	89.7	35	22	AAO12382	Human polypeptide
29	26	89.7	39	21	AAE58343	Lung cancer associ
30	26	89.7	39	22	AAO13753	Human polypeptide
31	26	89.7	40	22	AAO07978	Human polypeptide
32	26	89.7	42	22	AAU31189	Human polypeptide
33	26	89.7	42	22	AAO13336	Human polypeptide
34	26	89.7	43	18	AAW19897	Human polypeptide
35	26	89.7	43	22	AAO06431	Human polypeptide
36	26	89.7	43	22	AAO11676	Human polypeptide
37	26	89.7	43	22	AAO11697	Human polypeptide
38	26	89.7	43	22	AAO12990	Human polypeptide
39	26	89.7	44	22	AAU21548	Human novel foetal
40	26	89.7	45	22	AAO07116	Human polypeptide
41	26	89.7	46	22	AAO07664	Human polypeptide
42	26	89.7	46	22	AAO11996	Human polypeptide
43	26	89.7	47	22	AAO13529	Human polypeptide
44	26	89.7	47	22	AAO13774	Human polypeptide
45	26	89.7	48	22	AAW60604	Human reproductive

ALIGNMENTS

RESULT 1	AA870246	standard; peptide: 6 AA.
ID	AA870246	standard; peptide: 6 AA.
XX	AA870246	
AC	AA870246	
XX	04-MAY-2001	(first entry)
DE	Peptide for identifying fluorescent insensitive sites in GFP.	
XX		
KW	Affinity fluorescent protein; detection; reporter.	
XX		
OS	Synthetic.	
XX		
PN	WO200109177-A2.	
XX		
PD	08-FEB-2001.	
XX		
PE	28-JUL-2000; 2000WO-US20619.	
XX		
PR	29-JUL-1999; 99US-0146438.	
XX		
PA	(WHEED) WHITEHEAD INST BIOMEDICAL RES.	
XX		
PI	Matsudaire PT, Ehrlich DJ, Zhong Q, Freyson Y;	
XX		
DR	WPI: 2001-159852/16.	
XX		
PT	New affinity fluorescent protein comprising a modified fluorescent	
PT	protein having a heterologous amino acid sequence and a	
PT	ligand-activated protein binding site, for detecting target ligand in a	
PT	mixture of macromolecules or in a cell	
XX		
PS	Claim 12; Page 8; 44pp; English.	

XX The present invention relates to an affinity fluorescent protein
 CC (aFP) having a modified fluorescent protein molecule with a mutated
 CC fluorescent protein molecule and a heterologous amino acid sequence
 CC with a ligand-activated protein binding site. The invention is useful
 CC for detecting target ligands in a mixture of macromolecules or a cell,
 CC for detecting and monitoring a range of biological activities and as
 CC as a substitute for reporter-molecule labeled monoclonal or
 CC polyclonal antibodies.
 CC
 SQ Sequence 6 AA;
 SO
 QY 1 LEPRAS 6 100.0%; Score 29; DB 22; Length 6;
 |||||
 Db 1 lepras 6 Best Local Similarity 100.0%; Pred. No. 6,4e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 2
 AAU31081
 ID AAU31081 standard; Protein: 119 AA.
 XX
 AC AAU31081;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secreted protein #1572.
 XX
 KW Human: vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200179449-A2.
 XX
 PD 25-OCY-2001.
 XX
 PF 16-APR-2001; 2001WO-US08656.
 XX
 PR 18-APR-2000; 2000US-0552929.
 PR 26-JAN-2001; 2001US-0770160.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI: 2001-611725/70.
 XX
 PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -
 XX
 PS Claim 20; Page 400; 765pp; English.
 XX
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and

CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 CC
 SQ Sequence 119 AA;
 SO
 QY 1 LEPRAS 6 100.0%; Score 29; DB 22; Length 119;
 |||||
 Db 81 lepras 86 Best Local Similarity 100.0%; Pred. No. 88;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 3
 AAU39701
 ID AAU39701 standard; Protein: 125 AA.
 XX
 AC AAU39701;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #597.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 DR N-PSDB; AAS59508.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID NO 896; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 125 AA;

Query Match 100.0%; Score 29; DB 22; Length 125;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
| | | | |
DB 85 lepras 90

RESULT 4

AAW8386
ID AAW8386 standard; Protein; 256 AA.

XX
AC AAW8386;

XX 26-APR-1999 (first entry)

XX Mammalian Zneul polypeptide.

XX Zneul-1: neuro-growth factor-like protein; human; breast cancer;
KM glioblastoma; pituitary adenoma; Alzheimer's disease; therapy;
KM nerve regeneration; haematopoiesis; fertility; contraception;
XX antibody.

OS Class - Mammalia.

XX W09857983-A2.

XX 23-DEC-1998.

XX 18-JUN-1998; 98WO-US12763.

XX 18-JUN-1997; 97US-0878322.

XX 18-JUN-1997; 97US-0050143.

XX (ZYMO) ZYMOGENETICS INC.

XX Blumberg H, Jelinek LJ, Lehner JM, Sheppard PO;

XX Whitmore TE;

XX WPI; 1999-095324/08.

XX New mammalian Zneul polypeptides - used to, e.g. treat Alzheimer's
XX disease, cancer and to repopulate blood cells

XX Claim 6; Page 54; 70pp; English.

XX This polypeptide comprises a novel mammalian Zneul polypeptide.
XX Novel human Zneul (see AAW8381) is a new neuro-growth factor-like
XX protein that shows homology to Notch 4 and which may be involved in
XX EGF receptor pathways. Zneul can be used as a growth, maintenance,
XX or differentiation factor in the spinal cord, heart, spleen,
XX testis, thyroid and lymph nodes. It may also be used to treat
XX Alzheimer's disease, cancer, to repopulate blood cells after
XX chemotherapy, to stimulate myofibroblast proliferation, stimulate
XX or inhibit growth factors made in the placenta, in fertility and
XX contraception, or to regenerate nerves. Claimed Zneul polypeptides
XX (see also AAW8382-97), including specific domains of Zneul and
XX epitope-bearing portions of Zneul, can be used to raise specific
XX antibodies for use e.g. in diagnostic assays.

XX Sequence 256 AA;

Query Match 100.0%; Score 29; DB 20; Length 256;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
| | | | |
DB 146 lepras 151

RESULT 5

AAB93466
ID AAB93466 standard; Protein; 526 AA.

XX AAB93466;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:12736.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Claim 8; SEQ ID 12736; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.

XX Sequence 526 AA;

Query Match 100.0%; Score 29; DB 22; Length 526;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
|
Db 335 lepras 340

RESULT 6

ABG22566
ID ABB69129 standard; Protein: 635 AA.

AC ABB69129;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 34179.

KW Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI: 2001-656860/75.

DR N-PSDB: ABL13232.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -

PS Disclosure: SEQ ID NO 34179; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
sequences (ABL01840-ABL16175) and the encoded proteins
(ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 635 AA;

Query Match 100.0%; Score 29; DB 22; Length 635;

Best Local Similarity 100.0%; Pred. No. 4.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
|
Db 27 lepras 32

RESULT 7

ABG22566
ID ABB22566 standard; Protein: 637 AA.

XX ABB22566;
AC ABB22566;
DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #22557.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB: AAS86753.

PT New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity -

Claim 20; SEQ ID NO 52925; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
polypeptide (II) sequences. (I) is useful as hybridisation probes,
polymerase chain reaction (PCR) primers, oligomers, and for chromosome
and gene mapping, and in recombinant production of (II). The
polynucleotides are also used in diagnostics as expressed sequence tags
for identifying expressed genes. (I) is useful in gene therapy techniques
to restore normal activity of (II) or to treat disease states involving
(II). (II) is useful for generating antibodies against it, detecting or
quantitating a polypeptide in tissue, as molecular weight markers and as
a food supplement. (II) and its binding partners are useful in medical
imaging of sites expressing (II). (I) and (II) are useful for treating
disorders involving aberrant protein expression or biological activity.
The polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. ABB00010-ABG30377 represent novel human
diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 637 AA;

Query Match 100.0%; Score 29; DB 22; Length 637;

Best Local Similarity 100.0%; Pred. No. 4.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
|
Db 202 lepras 207

RESULT 8

AAW88383
ID AAW88383 standard; Protein: 708 AA.

AC AAM88383;
XX
DT 26-APR-1999 (first entry)
XX
DE Mammalian Zneul polypeptide.
XX
KW Zneul-1; neuro-growth factor-like protein; human; breast cancer;
KM glioblastoma; pituitary adenoma; Alzheimer's disease; therapy;
KW nerve regeneration; haematopoiesis; fertility; contraception;
KW antibody.
XX
OS Class - Mammalia.
XX
PN WO9857983-A2.
XX
PD 23-DEC-1998.
XX
PF 18-JUN-1998; 98WO-US12763.
XX
PR 18-JUN-1997; 97US-0878322.
PR 18-JUN-1997; 97US-0050143.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Blumberg H, Jelinek LJ, Lehner JM, Sheppard PO;
PI Whitmore TE;
XX
DR WPI: 1999-095324/08.
XX
PT New mammalian Zneul polypeptides - used to, e.g. treat Alzheimer's
PT disease, cancer and to repopulate blood cells
XX
PS Claim 6; Page 51-53; 70pp; English.
XX
EC This polypeptide comprises a novel mammalian Zneul polypeptide.
CC Novel human Zneul (see AAM88381) is a new neuro-growth factor-like
CC protein that shows homology to Notch 4 and which may be involved in
CC EGF receptor pathways. Zneul can be used as a growth, maintenance,
CC or differentiation factor in the spinal cord, heart, spleen,
EC testis, thyroid and lymph nodes. It may also be used to treat
CC Alzheimer's disease, cancer, to repopulate blood cells after
CC chemotherapy, to stimulate myoblast proliferation, stimulate
CC or inhibit growth factors made in the placenta, in fertility and
CC contraception, or to regenerate nerves. Claimed Zneul polypeptides
CC (see also AAM88382-97), including specific domains of Zneul and
CC epitope-bearing portions of Zneul, can be used to raise specific
CC antibodies for use e.g. in diagnostic assays.
CC
SQ Sequence 708 AA;

Query Match 100.0%; Score 29; DB 20; Length 708;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LEPRAS 6
Db 598 Lepras 603

RESULT 9
AAM38679
ID AAM38679 standard; Protein; 966 AA.
XX
AC AAM38679;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 1824.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR N-PSDB; AA157835.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 3; SEQ ID NO 1824; 10078pp; English.
XX
EC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
CC
SQ Sequence 966 AA;

Query Match 100.0%; Score 29; DB 22; Length 966;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LEPRAS 6
Db 147 Lepras 152

RESULT 10
AAM38678
ID AAM38678 standard; Protein; 1013 AA.
XX
AC AAM38678;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 1823.

XX	Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW	chemokineic; thrombolytic; drug screening; arthritis; inflammation;
KW	leukaemia.
OS	
XX	
XX	Homo sapiens.
PN	
XX	WO20015312-A1.
PD	
XX	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-US34263.
XX	
PR	21-JAN-2000; 2000US-0486725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0596042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX	
DR	WPI: 2001-442253/47.
DR	N-PSDB; AAI57834.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders
PT	such as central nervous system injuries -
XX	
PS	Example 3; SEQ ID NO 1823; 10078bp; English.
XX	
CC	The invention relates to human nucleic acids (AA157798-AA161369) and
CC	the encoded polypeptides (AAM38642-AAM42213) with noctropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilisation of the activities such as: Immune system suppression,
CC	Activin/inhibin activity, chemotactic/chemokineic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemia and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
XX	
XX	Sequence 1013 AA;
XX	
QY	1 LEPRAS 6
DB	147 Lepras 152
XX	
XX	Query Match 100.0%; Score 29; DB 22; Length 1013;
XX	Best Local Similarity 100.0%; Pred. No. 7.2e+02;
XX	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX	
XX	RESULT 11
XX	AAM38680
XX	ID AAM38680 standard; Protein: 1025 AA.
XX	
XX	AAM38680;

XX	22-OCT-2001	(first entry)
DT		
XX		
DE		
XX	Human polypeptide SEQ ID NO 1825.	
KW	Human; neotropic; immunosuppressant; cyostatic; gene therapy; cancer;	
KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KM	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KM	chemokineic; thrombolytic; drug screening; arthritis; inflammation;	
XX	Leukemia.	
XX		
OS	Homo sapiens.	
PN		
FN	WO200153312-A1.	
PD		
FD	26-JUL-2001.	
XX		
PE	26-DEC-2000; 2000WO-US34263.	
XX		
PR	21-JAN-2000; 2000US-0488725.	
PR	25-APR-2000; 2000US-0552317.	
PR	09-JUL-2000; 2000US-0598042.	
PR	19-JUL-2000; 2000US-0620312.	
PR	03-AUG-2000; 2000US-0653450.	
PR	14-SEP-2000; 2000US-0662191.	
PR	19-OCT-2000; 2000US-0693036.	
PR	29-NOV-2000; 2000US-0727344.	
XX		
PA	(HXSE-) HXSEQ INC.	
XX		
F1	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	
P1	Wang Z, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;	
P1	Zhao QA, Zhou P, Goodrich R, Drmanac RT;	
XX		
DR	WPI: 2001-442253/47.	
DR	N-PSDB: AA157836.	
PT		
PT	Novel nucleic acids and polypeptides, useful for treating disorders	
XX	such as central nervous system injuries -	
PS		
XX	Example 3; SEQ ID NO 1825; 10078pp; English.	
CC	The invention relates to human nucleic acids (AA157798-AA161369) and	
CC	the encoded polypeptides (AAM8642-AA442213) with neotropic,	
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful	
CC	in gene therapy. A composition containing a polypeptide or polynucleotide	
CC	of the invention may be used to treat diseases of the peripheral nervous	
CC	system, such as peripheral nervous injuries, peripheral neuropathy and	
CC	localised neuropathies and central nervous system diseases, such as	
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic	
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the	
CC	utilisation of the activities such as: Immune system suppression,	
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic	
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,	
CC	assays for receptor activity, arthritis and inflammation, leukemias and	
C	C.N.S disorders.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification.	
XX		
SQ	Sequence 1025 AA;	
Query Match	100.0%; Score 29; DB 22; Length 1025;	
Best Local Similarity	100.0%; Prid. No. 7.3e+02;	
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 LEPRS 6 	
Db	147 leptas 152	
RESULT 12		

AB61324
 ID ABB61324 standard; Protein; 1077 AA.
 XX
 AC ABB61324;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 10764.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 XX
 DR N-PSDB; ABL05427.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 10764; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB572072).
 CC
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 CC Sequence 1077 AA;
 SQ

Query Match 100.0%; Score 29; DB 22; Length 1077;
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
 |||||
 Db 828 lepras 833

RESULT 13
 AAY78795
 ID AAY78795 standard; Protein; 1219 AA.
 XX
 AC AAY78795;
 XX
 DT 19-MAY-2000 (first entry)
 XX
 DE Human antizual-2 (Az-2) amino acid sequence.
 XX
 KW Antizual-2; Az-2; human; breast cancer; malignancy progression marker;
 KW malignancy reversion; tumour suppressor.
 XX

OS Homo sapiens.
 XX
 PN WO200000503-A1.
 XX
 PD 06-JAN-2000.
 XX
 PF 25-JUN-1999; 99WO-US14482.
 XX
 PR 26-JUN-1998; 98US-0090747.
 XX
 PA (CHEN/) CHEN H.
 XX
 PA (BISS/) BISSELL M.
 XX
 PI Chen H, Bissell M;
 XX
 DR WPI: 2000-170903/15.
 XX
 DR N-PSDB; AA290112.
 XX
 PT New nucleic acid encoding a tumour suppressor or marker, used for
 PT diagnosis, monitoring progress or treatment, and gene therapy of breast
 PT cancer -
 XX
 PS Claim 6; Page 69-72; 120pp; English.
 XX
 CC This sequence represents the human antizual-2 (Az-2) amino acid
 CC sequence. The Az-2 gene is a variant of the antizual-1 (Az-1) gene (see
 CC AA290111). The Az-1 gene is located on chromosome 10q26, and encodes a
 CC protein that acts as a tumour suppressor or marker of malignancy
 CC progression or reversion. The Az-1 protein and its variants are tumour
 CC suppressors, Az-1 interacts with E-cadherin and beta-catenin. Detecting
 CC low levels of Az-1 nucleotide or amino acid sequences are used to
 CC diagnose a breast cell malignancy, also for monitoring disease
 CC progression, particularly assessment of therapeutic efficacy. The
 CC nucleotide sequence is used in vivo or ex vivo gene therapy, and Az-1
 CC polypeptides are used for treating or preventing breast cancer. Az-1
 CC polypeptides are also used to raise specific antibodies, for diagnostic
 CC detection of Az-1. Fragments of the Az-1 nucleotide sequence are useful
 CC as probes or primers for detecting expression of the Az-1 gene.
 CC
 CC Sequence 1219 AA;
 SQ

Query Match 100.0%; Score 29; DB 21; Length 1219;
 Best Local Similarity 100.0%; Pred. No. 8.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
 |||||
 Db 404 lepras 409

RESULT 14
 AAB66476
 ID AAB66476 standard; Protein; 2359 AA.
 XX
 AC AAB66476;
 XX
 DT 09-APR-2001 (first entry)
 XX
 DE Rat alpha-1H calcium channel protein.
 XX
 KW Rat; expressed sequence tag; EST; antiarrhythmic; anticonvulsant;
 KW hypotensive; cardiant; noctropic; T-type calcium channel subunit;
 KW cardiac hypertrophy; cardiac arrhythmia; hypertension; sleep disorder;
 KW epilepsy; alpha-1H calcium channel.
 XX
 OS Rattus sp.
 XX
 PN WO200102561-A2.
 XX
 PD 11-JAN-2001.
 XX
 PF 04-JUL-2000; 2000WO-CA00794.

XX 02-JUL-1999: 99US-0346794.
 XX (NEUR-) NEUROMED TECHNOLOGIES INC.
 XX Snutch TP, Baillie DL;
 XX WPI: 2001-123111/13.
 DR N-PSDB: AAF31678.
 XX Novel T-type calcium channel alpha-1 subunit gene useful for treating
 PT cardiac hypertrophy, cardiac arrhythmia, hypertension, sleep disorder
 PT and epilepsy
 XX
 PS Disclosure: Page 75-85; 103pp; English.
 XX
 CC The present sequence is given in a specification providing sequences
 CC and partial sequences for three types of mammalian (human and rat)
 CC T-type calcium channel subunits. An expression cassette has been
 CC generated which comprises a nucleotide sequence encoding a T-type
 CC calcium channel alpha_1 subunit operably linked to control sequences
 CC to effect its expression. The novel calcium channel nucleic acids and
 CC proteins are useful for treating conditions characterised by
 CC undesirable levels of T-type calcium channel activity such as cardiac
 CC hypertrophy, cardiac arrhythmia, hypertension, sleep disorder and
 CC epilepsy.
 XX
 SQ Sequence 2359 AA;
 Query Match 100.0%; Score 29; DB 22; Length 2359;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LEPRAS 6
 Db 1968 lepras 1973
 RESULT 15
 AAY05781
 ID AAY05781 standard: Protein: 2548 AA.
 XX
 AC AAY05781;
 XX
 DT 02-AUG-1999 (first entry)
 XX
 DE Human myosin IXa.
 XX
 KM Myosin IXa: human; Bardet-Biedl syndrome; Usher syndrome; diagnosis;
 KM therapy; hearing loss; deafness; retinitis pigmentosa; obesity;
 KM hypogonadism; sterility; polydactyly; brachydactyly; syndactyly;
 KM mental retardation; renal abnormality; kidney disease;
 KM hypertension; diabetes; cardiovascular abnormality.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..146
 FT /note= "N-terminal extension, specifically claimed
 FT in Claim 3b"
 FT Domain 147..719
 FT /note= "head domain, specifically claimed in Claim
 FT 3c"
 FT Binding-site 239..246
 FT /note= "ATP binding loop, specifically claimed in
 FT Claim 3d"
 FT 720..835
 FT /note= "minimal length 2nd insert, specifically
 FT claimed in Claim 3e"
 FT 972..1166
 FT /note= "light chain binding domain, specifically
 FT claimed in Claim 3f"

FT Domain 2074..2219
 FT /note= "GAP domain, specifically claimed in Claim
 FT 3g"
 XX W09919489-A1.
 XX
 PD 22-APR-1999.
 XX
 PF 14-OCT-1998; 98WO-US21971.
 XX
 PR 30-DEC-1997; 97US-0068953.
 PR 15-OCT-1997; 97US-0062858.
 PR 17-OCT-1997; 97US-0062241.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Adams A, Chiu CY, Duhl D, Gorman SW, Leng S, Sheffield V;
 PI Welch J;
 XX
 DR WPI: 1999-277643/23.
 DR N-PSDB: AAX25487.
 XX
 PT Myosin IXa and cyclic nucleotide gated channel-15 polypeptides
 PS Claim 3a; Fig 1; 69pp; English.
 XX
 CC The present sequence represents human myosin IXa. The sequence is
 CC predicted from a myosin IXa cDNA clone (see AAX25487) that was
 CC isolated in a search for novel transcripts in the human chromosomal
 CC region 15p22-23BAC which could be candidates for the Bardet-Biedl
 CC syndrome gene, which maps to this locus. Myosin IXa polypeptides
 CC comprising the present sequence, or residues 1-146, 147-719,
 CC 239-246, 720-835, 872-116 and 2074-2219 are claimed. The invention
 CC discloses myosin IXa and cyclic nucleotide gated channel-15
 CC (CNGC-15) polypeptides, polynucleotides, and compositions containing
 CC these polypeptides, polynucleotides, expression cassettes,
 CC transformed cells and antibodies. Conditions that can be treated
 CC by CNGC-15 and/or myosin IXa polypeptides, agonists and antagonists
 CC include Bardet-Biedl syndrome, Usher syndrome, hearing loss,
 CC retinitis pigmentosa, obesity, hypogonadism, sterility, polydactyly,
 CC brachydactyly, syndactyly, mental retardation, renal abnormalities,
 CC hypertension, diabetes and cardiovascular abnormalities (all
 CC claimed). Methods for the expression and detection of CNGC-15 and
 CC myosin IXa nucleotides and polypeptides are also provided.
 XX
 SQ Sequence 2548 AA;
 Query Match 100.0%; Score 29; DB 20; Length 2548;
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LEPRAS 6
 Db 2357 lepras 2362

Search completed: June 6, 2002, 13:05:28
 Job time: 100 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2002, 13:03:48 ; Search time 12.96 Seconds
(without alignments)
11.308 Million cell updates/sec

Title: US-09-627-383-1

Perfect score: 29

Sequence: 1 LEPRAS 6

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	1261	1 US-08-764-100-26	Sequence 26, Appl
2	29	100.0	1385	2 US-08-687-399-7	Sequence 7, Appl
3	29	100.0	2548	4 US-09-172-422-1	Sequence 1, Appl
4	29	100.0	2887	4 US-08-462-467B-2	Sequence 2, Appl
5	29	100.0	2887	4 US-08-462-467B-8	Sequence 8, Appl
6	29	100.0	3218	1 US-08-764-100-27	Sequence 27, Appl
7	26	89.7	43	2 US-08-652-816A-43	Sequence 43, Appl
8	26	89.7	1788	2 US-08-962-284-2	Sequence 2, Appl
9	26	89.7	7257	3 US-09-335-409-5	Sequence 5, Appl
10	26	89.7	7257	4 US-09-568-102-5	Sequence 5, Appl
11	26	89.7	7257	4 US-09-567-969-5	Sequence 5, Appl
12	26	89.7	7257	4 US-09-568-480-5	Sequence 5, Appl
13	26	89.7	7257	4 US-09-568-486-5	Sequence 5, Appl
14	26	89.7	7257	4 US-09-568-472-5	Sequence 5, Appl
15	25	88.2	102	2 US-08-808-982-8	Sequence 8, Appl
16	25	88.2	102	4 US-09-306-902A-8	Sequence 8, Appl
17	25	86.2	118	4 US-09-202-712-17	Sequence 17, Appl
18	25	86.2	159	3 US-08-390-353A-5	Sequence 6, Appl
19	25	86.2	221	3 US-08-390-353A-6	Sequence 6, Appl
20	25	86.2	297	1 US-07-866-560-6	Sequence 6, Appl
21	25	86.2	297	1 US-08-077-673-6	Sequence 6, Appl
22	25	86.2	297	1 US-08-478-992-6	Sequence 6, Appl
23	25	86.2	297	3 US-09-105-298-6	Sequence 6, Appl
24	25	86.2	297	3 US-08-706-281A-10	Sequence 10, Appl
25	25	86.2	297	4 US-09-097-231-10	Sequence 1, Appl
26	25	86.2	413	3 US-08-390-353A-7	Sequence 1, Appl
27	25	86.2	426	3 US-08-390-353A-1	Sequence 1, Appl

28	25	86.2	430	2 US-08-712-709-9	Sequence 9, Appl
29	25	86.2	430	3 US-09-111-444-9	Sequence 9, Appl
30	25	86.2	430	4 US-09-541-228-9	Sequence 9, Appl
31	25	86.2	431	2 US-08-712-709-5	Sequence 5, Appl
32	25	86.2	431	3 US-09-111-444-5	Sequence 5, Appl
33	25	86.2	431	3 US-09-541-228-5	Sequence 5, Appl
34	25	86.2	431	4 US-09-031-295-2	Sequence 7, Appl
35	25	86.2	599	2 US-08-954-333-7	Sequence 154, App
36	25	86.2	748	2 US-08-997-080-154	Sequence 154, App
37	25	86.2	748	2 US-08-997-362-154	Sequence 154, App
38	25	86.2	748	4 US-09-095-855-154	Sequence 154, App
39	25	86.2	748	4 US-09-324-542-154	Sequence 14, App
40	25	86.2	1399	4 US-08-462-467B-14	Sequence 4, Appl
41	25	86.2	1618	4 US-08-462-467B-4	Sequence 2, Appl
42	25	86.2	1711	2 US-08-342-930-2	Sequence 2, Appl
43	25	86.2	2588	3 US-08-936-135-2	Sequence 199, App
44	24	82.8	34	3 US-08-851-843A-199	Sequence 318, App
45	24	82.8	34	4 US-08-974-549A-318	

ALIGNMENTS

RESULT 1
US-08-764-100-26
Sequence 26, Application US/08764100
Patent No. 5773700
GENERAL INFORMATION:
APPLICANT: Van Ginsven J., Martinus Q.
APPLICANT: De Haan, Petrus T.
APPLICANT: Gielen L., Johannes J.
APPLICANT: Peters, Dirk
TITLE OF INVENTION: Goldbach, Robert W.
TITLE OF INVENTION: Improvements in or Relating to Organic
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sandoz Agro, Inc
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,100
FILING DATE: 06-DEC-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,064
FILING DATE:
APPLICATION NUMBER: US 08/032,235
FILING DATE: 17-MAR-1993
APPLICATION NUMBER: GB 9206016.9
FILING DATE: 19-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5773700f1s, Allen E.
REGISTRATION NUMBER: 34,490
REFERENCE/DOCKET NUMBER: 137-1061
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 354-3592
TELEFAX: (415) 857-1125
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 1261 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein

US-08-764-100-26

Query Match 100.0%; Score 29; DB 1; Length 1261;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
DB 800 LEPRAS 805

RESULT 2

US-08-687-399-7
Sequence 7, Application US/08687399
Patent No. 5928381
GENERAL INFORMATION:
APPLICANT: Toft, Annette H.
APPLICANT: Marcher, Dorthie
APPLICANT: Pedersen, Hanne H.
APPLICANT: Nilsson, Thomas E.
TITLE OF INVENTION: A Combined Desizing and Bleaching
TITLE OF INVENTION: Process
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: No. 5928381disk of No. 5928381th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,399
FILING DATE:
CLASSIFICATION: 008
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4127, 204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1385 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-687-399-7

Query Match 100.0%; Score 29; DB 2; Length 1385;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
DB 515 LEPRAS 520

RESULT 3

US-09-172-422-1
Sequence 1, Application US/09172422A
Patent No. 6300485
GENERAL INFORMATION:
APPLICANT: Adams, Arwen E.
APPLICANT: Chiu, Chai Ying

APPLICANT: Duhl, David
APPLICANT: Gorman, Susan W.
APPLICANT: Leng, Song
APPLICANT: Sheffield, Val
APPLICANT: Welch, Juliet
TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED
TITLE OF INVENTION: CHANNEL-15 (CNCG-15) POLYNUCLEOTIDES, POLYPEPTIDES,
TITLE OF INVENTION: COMPOSITIONS, METHODS, AND USES THEREOF
FILE REFERENCE: 200130.442
CURRENT APPLICATION NUMBER: US/09/172,422A
CURRENT FILING DATE: 1998-10-14
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 1
LENGTH: 2548
TYPE: PRT
ORGANISM: Homo sapien
US-09-172-422-1

Query Match 100.0%; Score 29; DB 4; Length 2548;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
DB 2357 LEPRAS 2362

RESULT 4

US-08-462-467B-2
Sequence 2, Application US/08462467B
Patent No. 6210899
GENERAL INFORMATION:
APPLICANT: Rosenbaum, Jan S
TITLE OF INVENTION: The use of a BMP Protein Receptor
TITLE OF INVENTION: Complex for Screening Bone Metabolism Actives and Cells
TITLE OF INVENTION: Co-Transfected With a Type II BMP Receptor and a Type I
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESSES:
ADDRESSEE: The Procter & Gamble Company
STREET: 11810 East Miami River Road
CITY: Ross
STATE: OH
COUNTRY: USA
ZIP: 45061
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,467B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hersko, Bart S.
REGISTRATION NUMBER: 32,572
REFERENCE/DOCKET NUMBER: 5474R
TELECOMMUNICATION INFORMATION:
TELEPHONE: (513) 627-0633
TELEFAX: (513) 627-0260
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2887 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-462-467B-2

Query Match 100.0%; Score 29; DB 4; Length 2887;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
Db 2034 LEPRAS 2039

RESULT 5
US-08-462-467B-8
; Sequence 8, Application US/08462467B
; Patent No. 6210899
; GENERAL INFORMATION:
; APPLICANT: Rosenbaum, Jan S
; TITLE OF INVENTION: The Use of a BMP Protein Receptor
; TITLE OF INVENTION: Complex for Screening Bone Metabolism Actives and Cells
; TITLE OF INVENTION: Co-Transfected With a Type II BMP Receptor and a Type I
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Procter & Gamble Company
; STREET: 11810 East Miami River Road
; CITY: Ross
; STATE: OH
; COUNTRY: USA
; ZIP: 45061
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,467B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hersko, Bart S.
; REGISTRATION NUMBER: 32,572
; REFERENCE/DOCKET NUMBER: 5474R
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (513) 627-0633
; TELEFAX: (513) 627-0260
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2887 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; US-08-462-467B-8

Query Match 100.0%; Score 29; DB 4; Length 2887;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
Db 2034 LEPRAS 2039

RESULT 6
US-08-764-100-27
; Sequence 27, Application US/08764100
; Patent No. 5773700
; GENERAL INFORMATION:
; APPLICANT: Van Grinsven J., Martinus Q.
; APPLICANT: De Haan, Petrus T.
; APPLICANT: Gieleen L., Johannes J.
; APPLICANT: Peters, Dirk
; APPLICANT: Goldbach, Robert W.
; TITLE OF INVENTION: Improvements in or Relating to Organic

TITLE OF INVENTION: Compounds
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sandoz Agro, Inc
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,100
FILING DATE: 06-DEC-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,064
FILING DATE:
APPLICATION NUMBER: US 08/032,235
FILING DATE: 17-MAR-1993
APPLICATION NUMBER: GB 9206016.9
FILING DATE: 19-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: NO. 57737001s, Allen E.
REGISTRATION NUMBER: 34,490
REFERENCE/DOCKET NUMBER: 137-1061
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 354-3592
TELEFAX: (415) 857-1125
INFORMATION FOR SRO ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 3218 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-764-100-27

Query Match 100.0%; Score 29; DB 1; Length 3218;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
Db 312 LEPRAS 317

RESULT 7
US-08-652-816A-43
; Sequence 43, Application US/08652816A
; Patent No. 5872215
; GENERAL INFORMATION:
; APPLICANT: Osbourn, JK
; APPLICANT: Allen, DJ
; APPLICANT: McCafferty, JG
; TITLE OF INVENTION: Specific binding members, materials and
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652.816A
; FILING DATE: 23-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9125579.4
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9125579.8
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206318.9
; FILING DATE: 24-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206372.6
; FILING DATE: 23-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9525004.9
; FILING DATE: 07-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610824.6
; FILING DATE: 23-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/02240
; FILING DATE: 02-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/244.597
; FILING DATE: 01-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: David W. Clough
; REGISTRATION NUMBER: 36.107
; REFERENCE/DOCKET NUMBER: 28111/33308
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-652-816A-43

Query Match      89.7%; Score 26; DB 2; Length 43;
Best Local Similarity 83.3%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
   11111
DB 32 LEPRAS 37

RESULT 8
US-08-962-284-2
; Sequence 2, Application US/08962284
; Patent No. 5985608
; GENERAL INFORMATION:
; APPLICANT: Luna, Elizabeth J.
; APPLICANT: Pestonjamas, Kerst N.
; APPLICANT: Pope, Robert K.
; APPLICANT: Wulfkuhle, Julia D.
; TITLE OF INVENTION: ACTIN-BINDING POLYPEPTIDES
; TITLE OF INVENTION: AND NUCLEIC ACIDS ENCODING THE SAME
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```

;
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/962.284
; FILING DATE: 31-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, Peter J.
; REGISTRATION NUMBER: 32.983
; REFERENCE/DOCKET NUMBER: 07917/058001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1788 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
; US-08-962-284-2

Query Match      89.7%; Score 26; DB 2; Length 1788;
Best Local Similarity 83.3%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
   11111
DB 1033 VEPRAS 1038

RESULT 9
US-09-335-409-5
; Sequence 5, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHLONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335.409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
; US-09-335-409-5

Query Match      89.7%; Score 26; DB 3; Length 7257;
Best Local Similarity 83.3%; Pred. No. 5.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
   11111
DB 2587 LDPRAS 2592

RESULT 10
US-09-568-102-5
; Sequence 5, Application US/09568102
; Patent No. 6346404
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
```

```
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,102
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-102-5
```

```
Query Match      89.7%; Score 26; DB 4; Length 7257;
Best Local Similarity 83.3%; Pred. No. 5.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 LEPRAS 6
        1:|||||
Db      2587 LDPRAS 2592
```

```
RESULT 11
US-09-567-969-5
; Sequence 5, Application US/09567969
; Patent No. 6355457
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/567,969
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-567-969-5
```

```
Query Match      89.7%; Score 26; DB 4; Length 7257;
Best Local Similarity 83.3%; Pred. No. 5.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 LEPRAS 6
        1:|||||
Db      2587 LDPRAS 2592
```

```
RESULT 12
US-09-568-480-5
; Sequence 5, Application US/09568480
; Patent No. 6355458
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
```

```
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,480
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-480-5
```

```
Query Match      89.7%; Score 26; DB 4; Length 7257;
Best Local Similarity 83.3%; Pred. No. 5.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 LEPRAS 6
        1:|||||
Db      2587 LDPRAS 2592
```

```
RESULT 13
US-09-568-486-5
; Sequence 5, Application US/09568486
; Patent No. 6355458
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,486
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-486-5
```

```
Query Match      89.7%; Score 26; DB 4; Length 7257;
Best Local Similarity 83.3%; Pred. No. 5.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 LEPRAS 6
        1:|||||
Db      2587 LDPRAS 2592
```

```
RESULT 14
US-09-568-472-5
; Sequence 5, Application US/09568472
; Patent No. 6358719
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,472
```

;; CURRENT FILING DATE: 2000-05-10
;; PRIOR APPLICATION NUMBER: 09/335,409
;; PRIOR FILING DATE: 1999-06-17
;; NUMBER OF SEQ ID NOS: 30
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 5
;; LENGTH: 7257
;; TYPE: PRT
;; ORGANISM: Sorangium cellulosum
US-09-568-472-5

Query Match 89.7%; Score 26; DB 4; Length 7257;
Best Local Similarity 83.3%; Pred. No. 5.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
|:||||
Db 2587 LDPRAS 2592

RESULT 15
US-08-808-982-8
; Sequence 8, Application US/08808982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-808-982-8

Query Match 86.2%; Score 25; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 5
|:||||
Db 13 LEPRAS 17

Search completed: June 6, 2002, 13:04:11
Job time: 23 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2002, 13:03:48 ; Search time 15 Seconds
(without alignments)
38.436 Million cell updates/sec

Title: US-09-627-383-1

Perfect score: 29

Sequence: 1 LEPRAS 6

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_71:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description
1	29	100.0	162 2	A75359
2	29	100.0	242 2	G82642
3	29	100.0	1397 2	T46354
4	29	100.0	2626 2	T31099
5	27	93.1	201 2	JC5476
6	27	93.1	366 2	G95376
7	27	93.1	388 2	AG3444
8	27	93.1	422 2	H82236
9	26	89.7	214 2	D73576
10	26	89.7	238 2	S50809
11	26	89.7	254 2	AE3355
12	26	89.7	321 2	F75287
13	26	89.7	352 2	G83636
14	26	89.7	391 2	T38882
15	26	89.7	461 2	AH0373
16	26	89.7	465 1	F64630
17	26	89.7	465 1	D71884
18	26	89.7	477 2	T39867
19	26	89.7	561 2	S62552
20	26	89.7	562 2	AG6713
21	26	89.7	601 2	T32486
22	26	89.7	609 2	G81750
23	26	89.7	710 1	S70965
24	26	89.7	821 1	B34488
25	26	89.7	836 2	T07322
26	26	89.7	986 2	T41809
27	26	89.7	987 2	T40241
28	26	89.7	988 2	JQ1477
29	26	89.7	988 4	S58114

30	26	89.7	1187 2	JE0347	hypothetical prote
31	26	89.7	1224 2	T25770	hypothetical prote
32	26	89.7	1622 2	T45240	hypothetical prote
33	26	89.7	2195 2	S61103	Sec16 protein - ye
34	26	89.7	4753 1	A47437	LDL-receptor-relat
35	25	86.2	109 2	T34639	hypothetical prote
36	25	86.2	114 2	C72759	hypothetical prote
37	25	86.2	115 2	G83426	hypothetical prote
38	25	86.2	125 2	H81997	probable lipoprote
39	25	86.2	125 2	F81226	lipoprotein, proba
40	25	86.2	132 2	T46940	hypothetical prote
41	25	86.2	166 2	AF2899	conserved hypotet
42	25	86.2	168 2	A84156	single-strand DNA-
43	25	86.2	172 2	G97674	hypothetical prote
44	25	86.2	208 2	B87468	conserved hypotet
45	25	86.2	228 2	A83305	hypothetical prote

ALIGNMENTS

```

RESULT 1
A75359
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 15-Sep-2000
C:Accession: A75359
R:Waller, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Yamathayan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioreistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: A75359
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-162 <NR1>
A:Cross-references: GB:AE002017; GB:AE000513; NID:g6459527; PIDN:AAF1308.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1748
A:Map position: 1
C:Superfamily: Deinococcus radiodurans hypothetical protein DR1748

Query Match 100.0%; Score 29; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
DB 151 LEPRAS 156

RESULT 2
G82642
conserved hypothetical protein XF1737 [Imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: G82642
A:Molecule type: DNA
A:Residues: 1-242 <SIM>
A:Cross-references: GB:AE003997; GB:AE003849; NID:g9106805; PIDN:AAF84546.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre

```

as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
 Submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kltajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laiz
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, F.
 A:Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF1737

Query Match 100.0%; Score 29; DB 2; Length 242;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
 Db 75 LEPRAS 80

RESULT 3
 T46354
 Hypothetical protein DKFZp434F1016.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
 C:Accession: T46354
 R:Kocher, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23037
 A:Accession: T46354
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1397 <AAA>
 A:Cross-references: EMBL:ALJ37287
 A:Experimental source: adult testis; clone DKFZp434F1016
 C:Genetics:
 A:Note: DKFZp434F1016.1
 A:Cross-family: protein kinase C zinc-binding repeat homology
 F:845-897/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match 100.0%; Score 29; DB 2; Length 1397;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
 Db 1206 LEPRAS 1211

RESULT 4
 T31099
 myosin-Rhocap protein, Myr 7 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 08-Sep-2000
 C:Accession: T31099
 R:Chiergatti, E.; Gaertner, A.; Scoeffler, H.E.; Baehler, M.
 submitted to the EMBL Data Library, October 1997
 A:Description: Myr 7 is a novel myosin-Rho-Gap molecule expressed in rat brain.
 A:Reference number: Z20982
 A:Accession: T31099
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-2626 <CHI>
 A:Cross-references: EMBL:AJ001713; NID:e1357083; PID:e1357084; PIDN:CA04946.1
 A:Experimental source: strain Sprague-Dawley
 C:Genetics:
 A:Gene: myo9a

C:Superfamily: protein kinase C zinc-binding repeat homology; myosin motor domain hom
 F:149-1005/Domain: myosin motor domain homology #status atypical <MMO>
 F:2068-2116/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match 100.0%; Score 29; DB 2; Length 2626;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
 Db 2443 LEPRAS 2448

RESULT 5
 JC5476
 Junction-specific DNA-binding protein ruva - Pseudomonas aeruginosa
 N:Alternate names: Holliday junction DNA helicase ruva
 C:Species: Pseudomonas aeruginosa
 C:Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 31-Dec-2000
 C:Accession: JC5476; H83524
 R:Hishida, T.; Iwasaki, H.; Ishioke, K.; Shinagawa, H.
 Gene 182, 63-70, 1996
 A>Title: Molecular analysis of the Pseudomonas aeruginosa genes, ruva, ruvA and ruvC,
 A:Reference number: JC5476; MUID:97136691
 A:Accession: JC5476
 A:Molecule type: DNA
 A:Residues: 1-201 <HIS>
 A:Cross-references: DBJ:D83138; NID:g1183837; PIDN:BA11818.1; PID:g1183841
 A:Experimental source: strain PA01
 A:Note: the authors translated the initiation codon GTG for residue 1 as Met
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; L
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
 A:Reference number: AB2950; MUID:20437337
 A:Accession: H83524
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-201 <STO>
 A:Cross-references: GB:AE004530; GB:AE004091; NID:g9946865; PIDN:AG04355.1; GSPDB:GN
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: ruva; PA0966
 A:Start codon: GTG
 C:Function:
 A:Description: acting as specificity factor that targets ruvB protein; ruvA protein a
 ate stages of homologous genetic recombination and DNA repair
 C:Superfamily: holliday junction DNA helicase ruva
 C:Keywords: DNA binding; DNA recombination; DNA repair
 F:106-126/Domain: helix-hairpin-helix #status predicted <HHN>

Query Match 93.1%; Score 27; DB 2; Length 201;
 Best Local Similarity 83.3%; Pred. No. 53;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
 Db 146 MEPRAS 151

RESULT 6
 G95376
 conserved hypothetical protein SMA1676 [imported] - Sinorhizobium meliloti (strain 10
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: G95376
 R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B
 .; Kalmun, S.; Keating, D.H.; Palm, C.; Beck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A>Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli

A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: G95376
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-366 <KUR>
A:Cross-references: GB:AE006469; PIDN:AAK65577.1; PID:g14524057; GSPDB:GN00165
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galbert, F.; Flann, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, M.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Sma1676
A:Genome: plasmid

Query Match 93.1%; Score 27; DB 2; Length 366;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
:|||||
DB 292 IEPRAS 297

RESULT 7
AG3444
transporter BME11541 [imported] - *Brucella melitensis* (strain 16M)
C:Species: *Brucella melitensis*
-C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AG3444
R:DeVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, .; Marut, M.; Goldsman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A:Reference number: AD3252; PMID:11756688
A:Accession: AG3444
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-388 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL52722.1; PID:g17983551; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11541
A:Map position: 1

Query Match 93.1%; Score 27; DB 2; Length 388;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
:|||||
DB 236 MEPRAS 241

RESULT 8
H82236
iron-containing alcohol dehydrogenase VC1147 [imported] - *Vibrio cholerae* (strain N16961)
C:Species: *Vibrio cholerae*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: H82236
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: AG2035; MUID:20406833

A:Accession: H82236
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-422 <HEI>
A:Cross-references: GB:AE004194; GB:AE003852; NID:g9655612; PIDN:AAF94306.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1147
A:Map position: 1
C:Superfamily: lactaldehyde reductase; lactaldehyde reductase homology

Query Match 93.1%; Score 27; DB 2; Length 422;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
:|||||
DB 396 IEPRAS 401

RESULT 9
D75576
oxidoreductase, iron-sulfur subunit - *Deinococcus radiodurans* (strain R1)
C:Species: *Deinococcus radiodurans*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: D75576
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
M.; Shen, M.; Vamathevan, J.C.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A:Reference number: A75250; MUID:20036896
A:Accession: D75576
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-214 <WHI>
A:Cross-references: GB:AE001863; GB:AE001825; NID:g6460670; PIDN:AAF12402.1; PID:g646
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0235
A:Map position: 2
C:Superfamily: isoguinoline 1-oxidoreductase alpha chain; ferredoxin [2Fe-2S] homolog
C:Keywords: 2Fe-2S; metalloprotein
F:68/73/76/88/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 89.7%; Score 26; DB 2; Length 214;
Best Local Similarity 83.3%; Pred. No. 99;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
:|||||
DB 45 LDPRAS 50

RESULT 10
S50809
ribosomal protein L17, mitochondrial - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: HRD238 protein; mitochondrial ribosomal protein yml8; protein J112
C:Species: *Saccharomyces cerevisiae*
C:Date: 13-Jan-1995 #sequence_revision 08-Sep-1995 #text_change 13-Aug-1999
C:Accession: S50809; SA7128; S14890; S19235; S56837
R:Vandenbol, M.; Durand, P.; Dion, C.; Portetelle, D.; Hilger, F.
Yeast 11, 57-60, 1995
A:Title: Sequence of a 17.1 kb DNA fragment from chromosome X of *Saccharomyces cerevi*
A:Reference number: S50798; MUID:95282514
A:Accession: S50809
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-238 <VAN>
A:Cross-references: EMBL:234288; NID:g498992; PIDN:CAA84060.1; PID:g499004
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994

R;Vandenbol, M.; Durand, P.; Dion, C.; Portetelle, D.; Hilger, F.
submitted to the EMBL Data Library, June 1994
A:Description: Sequence analysis of a 17.1 kb DNA fragment from chromosome X of *Saccharomyces cerevisiae*
A:Reference number: S47117
A:Accession: S47118
A:Molecule type: DNA
A:Residues: 1-238 <VAM>
A:Cross-references: EMBL:Z34288; NID:g498992; PIDN:CAA84060.1; PID:g499004
R;Kitakawa, M.; Grohmann, L.; Graack, H.R.; Isono, K.
Nucleic Acids Res. 18, 1521-1529, 1990
A:Title: Cloning and characterization of nuclear genes for two mitochondrial ribosomal proteins
A:Reference number: S14889; MUID:90221879
A:Accession: S14890
A:Molecule type: DNA
A:Residues: 1-81, 'G', 83-238 <KIT>
A:Cross-references: EMBL:X53841; NID:g3981; PIDN:CAA37834.1; PID:g3982
A:Accession: S19235
A:Molecule type: protein
A:Residues: 2-28 <KIT2>
A:Experimental source: Strain DC5, YNN27
R;Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56835
A:Accession: S56837
A:Molecule type: DNA
A:Residues: 1-238 <POH>
A:Cross-references: EMBL:Z49338; NID:g1008210; PIDN:CAA89354.1; PID:g1008211; MIPS:YJL06
C:Genetics:
A:Gene: SGD:MRPL8
A:Cross-references: SGD:S0003599; MIPS:YJL063C
A:Map position: 10L
A:Genome: nuclear
C:Superfamily: Escherichia coli ribosomal protein L17
C:Keywords: mitochondrion; protein biosynthesis; ribosome
F:2-238/Product: ribosomal protein L17 #status experimental <MAT>

Query Match 89.7%; Score 26; DB 2; Length 238;
Best Local Similarity 83.3%; Pred. NO. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
|||||
Db 111 LEPRAN 116

RESULT 11
AE3355
undecaprenyl pyrophosphate synthetase (EC 2.5.1.31) [imported] - *Brucella melitensis* (str. C)
C:Species: *Brucella melitensis*
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C:Accession: AE3355
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova, P.; Mazur, M.; Gotsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, A.; Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis* strain 16M
A:Reference number: AD3252; PMID:1175688
A:Accession: AE3355
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-254 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAJ52008.1; PID:g17982771; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0827
A:Map position: 1
C:Superfamily: conserved hypothetical protein YAR002C
C:Keywords: transferase

Query Match 89.7%; Score 26; DB 2; Length 254;
Best Local Similarity 83.3%; Pred. NO. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
|||||
Db 232 VEPRAS 237

RESULT 12
F75287
probable DNA polymerase III subunit - *Deinococcus radiodurans* (strain R1)
C:Species: *Deinococcus radiodurans*
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: F75287
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Shen, M.; Vamthyan, J.D.; Lam, P.; McDonald, L.; Ulterback, T.; Zalewski, C.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A:Reference number: A75250; MUID:20036896
A:Accession: F75287
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <WHI>
A:Cross-references: GB:AE002064; GB:AE000513; NID:g6460134; PIDN:AAFL1878.1; PID:g646
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2332
A:Map position: 1

Query Match 89.7%; Score 26; DB 2; Length 321;
Best Local Similarity 83.3%; Pred. NO. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
|||||
Db 86 LEPRAT 91

RESULT 13
G83636
conserved hypothetical protein PA0069 [imported] - *Pseudomonas aeruginosa* (strain PAO)
C:Species: *Pseudomonas aeruginosa*
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: G83636
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lior, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337
A:Accession: G83636
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-352 <STO>
A:Cross-references: GB:AE004446; GB:AE004091; NID:g9945886; PIDN:AAO3459.1; GSPDB:GN
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA0069

Query Match 89.7%; Score 26; DB 2; Length 352;
Best Local Similarity 83.3%; Pred. NO. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
|||||
Db 200 LEPRAN 205

RESULT 14
T38882
hypothetical protein SPAC4H3.02c - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 04-Mar-2000
 C:Accession: T38882
 R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
 submitted to the EMBL Data Library, February 1996
 A:Reference number: Z21807
 A:Accession: T38882
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-391 <MUR>
 A:Cross-references: EMBL:Z69380; PIDN:CA93341.1; GSPDB:GN00066; SPDB:SPAC4H3.02c
 A:Experimental source: strain 972h-; cosmid c4H3
 C:Genetics:
 A:Gene: SPDB:SPAC4H3.02c
 A:Map position: 1
 A:Introns: 56/2
 C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPAC4H3.02c

Query Match 89.7%; Score 26; DB 2; Length 391;
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 LEPRAS 6
 1:1111
 Db 37 LQPRAS 42

RESULT 15
 AH0373
 cysteine--trna ligase (EC 6.1.1.16) [Imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AH0373
 R:Parhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.;
 +deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 -A:Accession: AH0373
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-461 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC92315.1; PID:g15981026; GSPDB:GN00175
 C:Genetics:
 A:Gene: cyss
 C:Superfamily: cysteine--trna ligase
 C:Keywords: ligase

Query Match 89.7%; Score 26; DB 2; Length 461;
 Best Local Similarity 83.3%; Pred. No. 2.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 LEPRAS 6
 11111:
 Db 110 LEPRAT 115

Search completed: June 6, 2002, 13:04:33
 Job time: 45 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2002, 13:04:38 ; Search time 10.31 Seconds
(without alignments)
22.533 Million cell updates/sec

Title: US-09-627-383-1

Perfect score: 29

Sequence: 1 LEPRAS 6

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	1026	1 TACC2_HUMAN	O95359 homo sapien
2	27	93.1	201	1 RUVA_PSEAE	O51425 pseudomonas
3	26	89.7	238	1 RM08_YEAST	P22353 saccharomyc
4	26	89.7	335	1 B3G3_CRIGR	O9W477 cricetus
5	26	89.7	335	1 B3G3_HUMAN	O94766 homo sapien
6	26	89.7	335	1 B3G3_MOUSE	P58158 mus musculu
7	26	89.7	331	1 YAY2_SCHPO	O10210 schizosach
8	26	89.7	465	1 SYC_HELPJ	O9ZK46 helicobacte
9	26	89.7	465	1 SYC_HELPJ	P41259 helicobacte
10	26	89.7	561	1 YAI7_SCHPO	O09894 schizosach
11	26	89.7	591	1 ALU1_HUMAN	P59188 homo sapien
12	26	89.7	710	1 PKN6_MYXXA	P54728 myxococcus
13	26	89.7	821	1 CAN3_RAT	P16239 rattus norv
14	26	89.7	836	1 RP0C_CHLVU	P56300 chlorella v
15	26	89.7	978	1 PEX6_RAT	P54777 rattus norv
16	26	89.7	986	1 DPOL_NFYBM	P41712 bombyx mori
17	26	89.7	988	1 TNP9_ECOLI	P51565 escherichia
18	26	89.7	2194	1 SC16_YEAST	P48415 saccharomyc
19	26	89.7	4753	1 LRP_CAEEL	O04833 caenorhabdi
20	25	86.2	190	1 DSR6_HUMAN	P57055 homo sapien
21	25	86.2	287	1 UL24_ILTVT	P23966 infectio
22	25	86.2	295	1 ACTR_SHEEP	O91U77 ovis aries
23	25	86.2	297	1 ACTR_BOVIN	P34974 bos taurus
24	25	86.2	300	1 MOVF_AMYLE	P03555 alfalfa mos
25	25	86.2	300	1 MOVF_AMYMA	P05672 alfalfa mos
26	25	86.2	300	1 MOVF_AMYST	P03566 alfalfa mos
27	25	86.2	300	1 MOVF_AMYST	P24265 alfalfa mos
28	25	86.2	309	1 NUSG_STRGB	P52852 streptomyce
29	25	86.2	430	1 SGK_RAT	O06226 rattus norv
30	25	86.2	431	1 SGK_HUMAN	O00141 homo sapien
31	25	86.2	431	1 SGK_MOUSE	O9WV66 mus musculu
32	25	86.2	431	1 SGK_RABIT	O9X18 oryctolagus
33	25	86.2	452	1 F26_YEAST	P32604 saccharomyc

34	25	86.2	453	1 HCAE_ECOLI	O47139 escherichia
35	25	86.2	495	1 MLP2_DROME	O24400 drosophila
36	25	86.2	505	1 Y4XG_RHISN	P55699 rhizobium s
37	25	86.2	563	1 YIM0_YEAST	P40475 saccharomyc
38	25	86.2	567	1 GPV_MOUSE	O08742 mus musculu
39	25	86.2	568	1 RRPV_MOUSE5	P27565 sendai viru
40	25	86.2	568	1 RRPV_MOUSE6	P14251 sendai viru
41	25	86.2	568	1 RRPV_SENDF	P14252 sendai viru
42	25	86.2	568	1 RRPV_SENDF	P04859 sendai viru
43	25	86.2	568	1 RRPV_SENDF	P04860 sendai viru
44	25	86.2	570	1 E7FD_ACICA	P94132 acinetobact
45	25	86.2	591	1 AGP1_ASCSU	O17045 ascaris suu

ALIGNMENTS

RESULT	1	STANDARD	PRT	1026 AA.
TACC2_HUMAN				
AC	O95359; O9NZR5; O9NZ41;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Transforming acidic coiled-coil-containing protein 2 (Anti Zua1-1) (AZU-1).			
GN	TACC2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=fetal brain, Brain, and Skeletal muscle;			
RX	MEDLINE=20570483; PubMed=11121038;			
RA	Gergely F., Karlsson C., Still I.H., Cowell J.K., Kilmartin J.,			
RA	Raff J.W.;			
RT	"The TACC domain identifies a family of centrosomal proteins that can			
RT	interact with microtubules."			
RL	Proc. Natl. Acad. Sci. U.S.A. 97:14352-14357(2000).			
RN	[2]			
RP	SEQUENCE OF 375-1026 FROM N.A. (ISOFORM 2).			
RX	MEDLINE=20214826; PubMed=10749935;			
RA	Chen H.-M., Schmeichel K.L., Mian I.S., Lelievre S., Petersen O.W.,			
RA	Bissell M.J.;			
RT	"AZU-1: a candidate breast tumor suppressor and biomarker for tumor			
RT	progression."			
RL	Mol. Biol. Cell 11:1357-1367(2000).			
RN	[3]			
RP	SEQUENCE OF 44-1026 FROM N.A. (ISOFORM 3).			
RA	Pu J., Li C., Rodriguez M., Banerjee D.;			
RT	"Expression of TACC2 protein mRNA in human microvascular endothelial			
RT	cells."			
RL	Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: MAY PLAY A ROLE IN ORGANIZING CENTROSOMAL MICROTUBULES.			
CC	MAY ACT AS A TUMOR SUPPRESSOR PROTEIN. MAY REPRESENT A TUMOR			
CC	PROGRESSION MARKER.			
CC	-1- SUBUNIT: INTERACTS WITH MICROTUBULES.			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR; CONCENTRATED AT CENTROSOMES.			
CC	-1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1 (SHOWN HERE), 2 AND 3; MAY			
CC	BE PRODUCED BY ALTERNATIVE SPLICING.			
CC	-1- SIMILARITY: BELONGS TO THE TACC FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF095791; AAC64968.2;			
DR	EMBL; AF176646; AAF63433.1; ALT_INIT.			

DR EMBL: AF220152; AAF29537.2; -
KW MIM: 605302; Nuclear protein; Alternative splicing.
FT DOMAIN 43 51 POLY-PRO.
FT DOMAIN 498 501 POLY-LYS.
FT DOMAIN 753 781 COILED COIL (POTENTIAL).
FT DOMAIN 824 1025 COILED COIL (POTENTIAL).
FT VARSPLIC 507 510 MISSING (IN ISOFORM 2).
FT VARSPLIC 711 787 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT CONFLICT 44 67 PAPPPPEVPEVSTQPPPEE -> MFWYKRGRGDMRP
FT VSIDGVVCS (IN REF. 3).
SO SEQUENCE 1026 AA; 112110 MW; E2575FCB446E9CF8 CRC64;

Query Match 100.0%; Score 29; DB 1; Length 1026;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
|11111|
Db 130 LEPRAS 135

RESULT 2
ID RUVA_PSEAE STANDARD: PRT: 201 AA.
AC 051425;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Holliday junction DNA helicase ruva.
GN RUVA OR PA0966.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
|1|
RN SEQUENCE FROM N.A.
RP STRAIN-ATCC 15692 / PA01;
RC MEDLINE=97136691; PubMed=8982068;
RA Hishida T., Iwasaki H., Ishiohara K., Shinagawa H.;
RT "Molecular analysis of the Pseudomonas aeruginosa genes, ruva, ruvB
RT and ruvC, involved in processing of homologous recombination
RT intermediates.";
RL Gene 182:63-70(1996).
|2|
RN SEQUENCE FROM N.A.
RP STRAIN-ATCC 15692 / PA01;
RC MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Gaidy R.L., Goltz L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Ladwig K., Lam R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
|1-|
CC FUNCTION: THE RUVA-RUVB COMPLEX IN THE PRESENCE OF ATP REMOVES
CC CRUCIFORM STRUCTURE IN SUPERCOILED DNA WITH PALINDROMIC SEQUENCE,
CC INDICATING THAT IT MAY PROMOTE STRAND EXCHANGE REACTIONS IN
CC HOMOLOGOUS RECOMBINATION. RUVA IS AN HELICASE THAT MEDIATES THE
CC HOLLIDAY JUNCTION MIGRATION BY LOCALIZED DENATURATION AND
CC RANNEHLING. RUVA STIMULATES, IN THE PRESENCE OF DNA, THE WEAK
CC APPASE ACTIVITY OF RUVB (BY SIMILARITY).
CC -1- SUBUNIT: FORMS A COMPLEX WITH RUVB.
CC -1- SIMILARITY: BELONGS TO THE RUVA FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D83138; BAA11818.1; -
DR EMBL: AE004530; AAG04355.1; -
DR HSSP: P08576; ICRK.
DR InterPro: IPR003583; HHH_1.
DR InterPro: IPR000085; RUVA.
DR Pfam: PF01330; RUVA_1.
DR Pfam: PF02904; RUVA_II; 1.
DR ProDom: PD006268; RUVA; 1.
DR SMART: SM00278; Hbh1; 2.
DR DNA repair: SOS response: ATP-binding; DNA recombination; Helicase;
KW Complete proteome.
SO SEQUENCE 201 AA; 21959 MW; 8EB3099077F87F23 CRC64;

Query Match 93.1%; Score 27; DB 1; Length 201;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
|11111|
Db 146 MEPRAS 151

RESULT 3
ID RM08_YEAST STANDARD: PRT: 238 AA.
AC P22353;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Mitochondrial 60S ribosomal protein L8 (Yml8).
GN MRPL8 OR YJL063C OR J1125 OR HRD238.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
|1|
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP STRAIN-07173;
RC MEDLINE=90221879; PubMed=21813197;
RA Kitakawa M., Grohmann L., Graack H.-R., Isono K.;
RT "Cloning and characterization of nuclear genes for two mitochondrial
RT ribosomal proteins in Saccharomyces cerevisiae.";
RL Nucleic Acids Res. 18:1521-1529(1990).
|2|
RN SEQUENCE FROM N.A.
RP STRAIN-S288C.
RC MEDLINE=95282514; PubMed=7762302;
RA Vandenbol M., Durand P., Dion C., Portetelle D., Hilger F.;
RT "Sequence of a 17.1 kb DNA fragment from chromosome X of
RT Saccharomyces cerevisiae includes the mitochondrial ribosomal protein
RT L8.";
RL Yeast 11:57-60(1995).
|1-|
CC SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: BELONGS TO THE L17P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X53841; CAA37834.1; -
DR EMBL: Z34288; CAA84060.1; -
DR EMBL: Z49338; CAA89354.1; -
DR PIR: S14890; S14890.
DR PIR: S47128; S47128.
DR SGD: S0003599; MRPL8.

RT region of proteoglycans".
RN FEBS Lett. 459:415-420(1999).
RC X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 76-335.
RA TISSUE-LIVER:
RX MEDLINE-20507909; PubMed-10946001;
RA Pedersen L.C., Tsuchida K., Kitagawa H., Sugahara K., Darden T.A.,
Negishi M.;
RT Heparan/chondroitin sulfate biosynthesis. Structure and mechanism of
human glucuronyltransferase 1.";
RL J. Biol. Chem. 275:34580-34585(2000).
CC -1- FUNCTION: GLYCOSAMINOGLYCANS BIOSYNTHESIS. INVOLVED IN FORMING THE
LINKAGE TETRASACCHARIDE PRESENT IN HEPARAN SULFATE AND CHONDROITIN
SULFATE. TRANSFERS A GLUCURONIC ACID MOIETY FROM THE URIDINE
DIPHOSPHATE-GLUCURONIC ACID (UDP-GLCUA) TO THE COMMON LINKAGE
REGION TRISACCHARIDE GAL BETA 1-3GAL BETA 1-4XYL COVALENTLY BOUND
TO A SER RESIDUE AT THE GLYCOSAMINOGLYCAN ATTACHMENT SITE OF
PROTEOGLYCANS. CAN ALSO PLAY A ROLE IN THE BIOSYNTHESIS OF I2/HNK-
1 CARBOHYDRATE EPIPOE ON GLYCOPROTEINS. SHOWS STRICT SPECIFICITY
FOR GAL BETA1-3GAL BETA1-4XYL, EXHIBITING NEGLIGIBLE INCORPORATION
INTO OTHER GALACTOSIDE SUBSTRATES INCLUDING GALBETA1-3GAL BETA1-O-
BENZYL, GALBETA1-4GALNA6 AND GALBETA1-4GIC.
CC CATALYTIC ACTIVITY: UDP-glucuronate + 3-beta-D-galactosyl-4-
beta-D-galactosyl-O-beta-D-xylosylprotein -> UDP + 3-beta-D-
glucuronosyl-3-beta-D-galactosyl-4-beta-D-galactosyl-O-beta-D-
xylosylprotein.
CC COFACTOR: MANGANESE.
CC -1- ENZYME REGULATION: INHIBITED BY EDTA.
CC -1- PATHWAY: GLYCOSYLATION.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi.
CC -1- TISSUE SPECIFICITY: UBIQUITOUS (BUT WEAKLY EXPRESSED IN ALL TISSUES
EXAMINED).
CC -1- PFM: N-GLYCOSYLATED.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 43.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC
CC EMBL; AB009598; BAA34537.1; -
DR EMBL; AJ005865; CAA06742.1; -
DR MIM; 606374; -
DR PDB; 1FGG; 31-JAN-01.
KM Transferase; Glycoprotein; Transmembrane; Signal-anchor; Golgi stack;
KW Manganese; Multigene family; 3D-structure.
FT DOMAIN 1 7
FT TRANSMEM 8 28
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT LUMENAL, CATALYTIC (POTENTIAL).
FT INTERCHAIN.
FT ACT_SITE 300 300
FT ACT_SITE 281 281
FT METAL 196 196
FT METAL 33 33
FT MANGANESE.
FT MANGANESE.
FT MUTAGEN 301 301
FT C->A: LOSS OF DIMER FORMATION AND REDUCED
ACTIVITY.
FT C->A: ENZYME INACTIVATION AND LOSS OF
GLYCOSYLATION.
FT CONFLICT 204 204
FT SEQUENCE 335 AA; 37061 MW; 5ED450428636C0F CRC64;
Query Match 89.7%; Score 26; DB 1; Length 335;
Best Local Similarity 83.3%; Pred. No. 74;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 6
B3G3_MOUSE STANDARD; PRT; 335 AA.
ID B3G3_MOUSE
AC P58158;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase 3
DE (EC 2.4.1.135) (Beta-1,3-glucuronyltransferase 3)
DE (Glucuronosyltransferase-1) (GlcAT-1) (UDP-GLCUA:gal Beta-1,3-gal-R
DE glucuronyltransferase) (GLCUAT-1).
GN B3GAT3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Breast tumor;
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: GLYCOSAMINOGLYCANS BIOSYNTHESIS. INVOLVED IN FORMING THE
LINKAGE TETRASACCHARIDE PRESENT IN HEPARAN SULFATE AND CHONDROITIN
SULFATE. TRANSFERS A GLUCURONIC ACID MOIETY FROM THE URIDINE
DIPHOSPHATE-GLUCURONIC ACID (UDP-GLCUA) TO THE COMMON LINKAGE
REGION TRISACCHARIDE GAL BETA 1-3GAL BETA 1-4XYL COVALENTLY BOUND
TO A SER RESIDUE AT THE GLYCOSAMINOGLYCAN ATTACHMENT SITE OF
PROTEOGLYCANS. CAN ALSO PLAY A ROLE IN THE BIOSYNTHESIS OF I2/HNK-
1 CARBOHYDRATE EPIPOE ON GLYCOPROTEINS (BY SIMILARITY).
CC CATALYTIC ACTIVITY: UDP-glucuronate + 3-beta-D-galactosyl-4-
beta-D-galactosyl-O-beta-D-xylosylprotein -> UDP + 3-beta-D-
glucuronosyl-3-beta-D-galactosyl-4-beta-D-galactosyl-O-beta-D-
xylosylprotein.
CC COFACTOR: MANGANESE (BY SIMILARITY).
CC -1- PATHWAY: GLYCOSYLATION.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (By
similarity).
CC -1- PFM: N-GLYCOSYLATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 43.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC
CC EMBL; BC002103; AA02103.1; -
DR EMBL; BC002103; AA02103.1; -
KM Transferase; Glycoprotein; Transmembrane; Signal-anchor; Golgi stack;
KW Manganese; Multigene family.
FT DOMAIN 1 7
FT TRANSMEM 8 28
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT LUMENAL, CATALYTIC (POTENTIAL).
FT ACT_SITE 281 281
FT ACT_SITE 196 196
FT METAL 33 33
FT METAL 300 300
FT MANGANESE (BY SIMILARITY).
FT INTERCHAIN (BY SIMILARITY).
FT CONFLICT 300 300
FT SEQUENCE 335 AA; 37067 MW; 906ED2AB672F1EC CRC64;
Query Match 89.7%; Score 26; DB 1; Length 335;
Best Local Similarity 83.3%; Pred. No. 74;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
RESULT 7
YAY2_SCHPO STANDARD: PRT: 391 AA.
ID YAY2_SCHPO
AC Q10210:
DR 01-OCT-1996 (Rel. 34, Created)
DR 01-OCT-1996 (Rel. 34, Last sequence update)
DR 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 44.7 kDa protein C4H3.02C in chromosome I.
GN SPAC4H3.02C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RL Murphy L., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DDSI databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 269380; CAA93341.1;
DR Hypothetical protein.
KW SEQUENCE 391 AA; 44716 MW; 72D8F088B15E50A8 CRC64;
SQ
```

```
Query Match 89.7%; Score 26; DB 1; Length 391;
Best Local Similarity 83.3%; Pred. No. 87;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 LEPKAS 6
Db 37 LQPKAS 42
```

```
RESULT 8
SYN_HELPJ STANDARD: PRT: 465 AA.
AC 09ZKW6;
DR 30-MAY-2000 (Rel. 39, Created)
DR 30-MAY-2000 (Rel. 39, Last sequence update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE CysteinyI-trNA synthetase (EC 6.1.1.16) (Cysteine--trNA ligase)
DE (CysRS).
GN CYS5 OR JHP0818.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonghe B.L., Carmel G.,
RA Tummino P.J., Carnuso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -1- CATALYTIC ACTIVITY: ATP + L-cysteine + tRNA(Cys) = AMP +
CC diphosphate + L-cysteinyI-trNA(Cys).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
```

```
CC . STRONG, TO METHIONYL-TRNA SYNTHETASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001511; AAD06399.1;
DR InterPro: IPR002308; trNA-synt_1e.
DR InterPro: IPR001412; trNA-synt_1.
DR Pfam: PF01406; trNA-synt_1e; 1.
DR PRINTS: PR00983; TRNASYNTHCTS.
DR PROSITE: PS00178; AA-TRNA_LIGASE_I. FALSE_NEG.
KW Aminoacyl-trNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 29 39 "HIGH" REGION.
FT SITE 269 273 "KMSKS" REGION.
FT BINDING 272 272 ATP (BY SIMILARITY).
SQ SEQUENCE 465 AA; 53306 MW; 336AE8BAED04476A CRC64;
```

```
Query Match 89.7%; Score 26; DB 1; Length 465;
Best Local Similarity 83.3%; Pred. No. 11e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 LEPKAS 6
Db 109 LEPKAS 114
```

```
RESULT 9
SYN_HELPJ STANDARD: PRT: 465 AA.
AC P41259;
DR 01-FEB-1995 (Rel. 31, Created)
DR 01-NOV-1997 (Rel. 35, Last sequence update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE CysteinyI-trNA synthetase (EC 6.1.1.16) (Cysteine--trNA ligase)
DE (CysRS).
GN CYS5 OR HP0886.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
RN (2)
RP SEQUENCE OF 278-465 FROM N.A.
RC STRAIN=ATCC 49503 / 60190;
RX MEDLINE=94193753; PubMed=8144644;
RA Cover T.L., Tumuru M.K., Cao P., Thompson S.A., Blaser M.J.;
RT "Divergence of genetic sequences for the vacuolating cytotoxin among
RT Helicobacter pylori strains.";
RL J. Biol. Chem. 269:10566-10573(1994).
RN (3)
RP SEQUENCE OF 407-465 FROM N.A.
```

```

RC STRAIN-CCUG.17874 / NCNC.11638;
RX MEDLINE-94222514; Pubmed-8168917;
RA Phadnis S.H., Ilver D.J., Janzon L., Normark S., Westblom T.U.;
RT "Pathological significance and molecular characterization of the
RT vacuolating toxin gene of Helicobacter pylori.";
RL Infect. Immun. 62:1557-1565(1994).
CC -1- CATALYTIC ACTIVITY: ATP + L-cysteine + tRNA(Cys) = AMP +
CC diphosphate + L-cysteinyl-tRNA(Cys).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC STRONG, TO METHIONYL-TRNA SYNTHETASE.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AE000598; AAD07934.1; -.
CC EMBL; U05676; AAI17656.1; -.
CC EMBL; U07145; AAI18866.1; -.
CC PIR; C53739; C53739.
CC TIGR; HP0886; -.
CC InterPro; IPR002308; tRNA-synt_1e.
CC InterPro; IPR001412; tRNA-synt_1.
CC Pfam; PF01406; tRNA-synt_1e; 1.
CC PRINTS; PR00983; TRNASYNTHCS.
CC PROSITE; PS00178; AA-TRNA-LIGASE_I; FALSE NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 29 "HIGH" REGION.
FT BINDING 269 273 "KMSKS" REGION.
FT BINDING 272 272 ATP (BY SIMILARITY).
FT VARIANT 280 280 V -> I (IN STRAIN ATCC 49503).
FT VARIANT 332 332 T -> N (IN STRAIN ATCC 49503).
FT VARIANT 392 392 I -> V (IN STRAIN ATCC 49503).
FT VARIANT 432 432 R -> Q (IN STRAINS ATCC 49503 AND NCNC
11638).
FT VARIANT 434 434 D -> N (IN STRAIN NCNC 11638).
FT VARIANT 440 440 S -> H (IN STRAIN ATCC 49503).
SQ SEQUENCE 465 AA; 53131 MW; B7053D5B8CDB7F30 CRC64;
-----
Query Match 89.7%; Score 26; DB 1; Length 465;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
-----
QY 1 LEPRAS 6
Db 109 LEPRAS 114
-----
RESULT 10
YAI7_SCHPO STANDARD; PRT; 561 AA.
AC Q09894;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 62.5 kDa protein C24B11.07C in chromosome I.
GN SPAC24B11.07C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCB1_TaxID-4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA Odell C., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;

```

```

RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO S.POMBE SPEC1902.02.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; Z67757; CAA91772.1; -.
KW Hypothetical protein.
SQ SEQUENCE 561 AA; 62463 MW; 146D40D7CF42A096 CRC64;
-----
Query Match 89.7%; Score 26; DB 1; Length 561;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
-----
QY 1 LEPRAS 6
Db 527 VEPRAS 532
-----
RESULT 11
ALU1_HUMAN STANDARD; PRT; 591 AA.
AC P39188;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alu subfamily J sequence contamination warning entry.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCB1_TaxID-9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95021758; Pubmed-7935834;
RA Claverie J.-M., Makalowski W.;
RT "Alu alert.";
RL Nature 371:752-752(1994).
RN [2]
RP CONCEPT.
RX MEDLINE-92241891; Pubmed-1572661;
RA Claverie J.-M.;
RT "Identifying coding exons by similarity search: alu-derived and other
potentially misleading protein sequences.";
RL Genomics 12:838-841(1992).
RN [3]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE-88333009; Pubmed-3138422;
RA Quentin Y.;
RT "The Alu family developed through successive waves of fixation
closely connected with primate lineage history.";
RL J. Mol. Evol. 27:194-202(1988).
RN [4]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE-91178815; Pubmed-1706781;
RA Jurka J., Milosavljevic A.;
RT "Reconstruction and analysis of human Alu genes.";
RL J. Mol. Evol. 32:105-121(1991).
CC -1- MISCELLANEOUS: VARIOUS ANALYSES (SEE REF.3 AND REF.4) INDICATE
CC THAT ALU REPEATS FALL INTO 8 SUBFAMILIES. THEREFORE, 8 ALU WARNING
CC CONSENSUS SEQUENCES HAVE BEEN CONSTITUTED THAT CONTAIN ALL SIX
CC FRAMES CONCEPTUAL TRANSLATIONS OF EACH OF THESE CLASSES OF ALU
CC REPEATS.
CC -1- MISCELLANEOUS: ISOLATED 'X' INDICATES THE PRESENCE OF A STOP
CC CODON, 'XXX' IS USED TO SEPARATE THE VARIOUS TRANSLATION PHASES.
CC -1- CAUTION: THIS ALU ENTRY IS PROVIDED IN ORDER TO AVOID THE FURTHER
CC POLUTION OF PROTEIN SEQUENCE DATABASES WITH ALU-DERIVED AMINO
CC ACID SEQUENCES.

```

CC -1- CAUTION: ALU REPETITIVE SEQUENCES ARE INTERSPERSED IN HUMAN AND
CC PRIMATE GENOMES WITH AN AVERAGE SPACING OF 4 KB. SOME OF THEM ARE
CC ACTIVELY TRANSCRIBED BY POL. III. NORMAL TRANSCRIPTS MAY CONTAIN
CC ALU-DERIVED SEQUENCES IN 5' OR 3' UNTRANSLATED REGIONS. HOWEVER,
CC CDNA LIBRARIES ALSO CONTAIN PARTIAL AND/OR REARRANGED CDNAS
CC LOCATED WITH ALU-DERIVED SEQUENCE IN ANY ORIENTATION. ALTHOUGH ALU
CC ELEMENTS (ESPECIALLY SITUATED ON THE COMPLEMENTARY STRAND) HAVE A
CC GREAT POTENTIAL TO CREATE ADDITIONAL/ALTERNATIVE EXONS,
CC CONSIDERATION SHOULD BE GIVEN TO THE POSSIBILITY THAT THE PRESENCE
CC OF AN ALU IN AN OPEN READING FRAME MAY HAVE RESULTED FROM A
CC CLONING ARTIFACT OR MAY BE DUE TO MISINTERPRETATION OF SEQUENCING
CC DATA. THIS POINT HAS BEEN OVERLOOKED ON SEVERAL OCCASIONS, WITH
CC THE CONSEQUENCE OF ERRONEOUS ALU-DERIVED AMINO ACID SEQUENCES
CC BEING REPORTED.
CC -1- CAUTION: ANY SIGNIFICANT SIMILARITY OF A POTATIVE PROTEIN SEQUENCE
CC WITH AN ALU-TRANSLATED ENTRY MUST BE TAKEN AS A WARNING THAT A
CC PART OF ALU REPEAT MAY HAVE BEEN ARTIFACTUALLY INCLUDED IN THE
CC CODING NUCLEOTIDE SEQUENCE.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U14567; -, NOT_ANNOTATED_CD5.
CC DR Hypothetical protein.
CC KM 1 96 FRAME-1.
CC FT DOMAIN 100 195 FRAME-2.
CC FT DOMAIN 199 294 FRAME-3.
CC FT DOMAIN 298 393 FRAME-4.
CC FT DOMAIN 397 492 FRAME-5.
CC FT DOMAIN 496 591 FRAME-6.
CC SQ SEQUENCE 591 AA; 63790 MW; 665D395735519D95 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 591;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
|11111|
DB 22 LEPRSS 27

RESULT 12
PKN6_MYXXA STANDARD; PRT; 710 AA.
ID PKN6_MYXXA
AC P54738;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Serine/threonine-protein kinase pkn6 (EC 2.7.1.-).
GN PKN6.
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxococcales; Cystobacteriineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DZFI;
RX MEDLINE=96610380; PubMed=8733241;
RA Zhang W., Inouye M., Inouye S.-I.
RT "Reciprocal regulation of the differentiation of Myxococcus xanthus
RT by Pkn5 and Pkn6, eukaryotic-like Ser/Thr protein kinases.";
RL Mol. Microbiol. 20:435-447(1996).
CC -1- FUNCTION: PKN5 AND PKN6 MAY HAVE RECIPROCAL ROLES IN GROWTH AND
CC DEVELOPMENT. PKN6 MAY BE A TRANSMEMBRANE SENSOR OF EXTERNAL
CC SIGNALS FOR DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED CONSTITUTIVELY THROUGHOUT THE LIFE

CC CYCLE, WITH SLIGHT INCREASES AT AN EARLY STAGE OF DEVELOPMENT.
CC -1- PTM: AUTOPHOSPHORYLATED AT SERINE AND THREONINE RESIDUES.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U40556; AAB40050.1; -.
CC DR HSSP: P00523; 2PTK.
CC DR InterPro: IPR000719; Euk_pkinase.
CC DR InterPro: IPR002290; Ser_thr_pkinase.
CC DR Pfam: PF00069; pkinase; 1.
CC DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
CC DR PROSITE: PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
CC DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC KW Transferase; Serine/threonine-protein kinase; ATP-binding;
CC Transmembrane; Phosphorylation; Repeat.
CC FT DOMAIN 1 482 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 483 503 POTENTIAL.
CC FT DOMAIN 504 710 PERIPLASMIC (POTENTIAL).
CC FT DOMAIN 6 280 PROTEIN KINASE.
CC FT NP_BIND 12 20 ATP (BY SIMILARITY).
CC FT BINDING 38 38 ATP (BY SIMILARITY).
CC FT ACT_SITE 143 143 BY SIMILARITY.
CC FT DOMAIN 337 369 2 X 12 AA REPEAT OF P-V/A-D-S-T-S-P-T'-T'-
CC FT REPEAT 337 348 P-N-P.
CC FT REPEAT 358 369 1.
CC FT SEQUENCE 710 AA; 74621 MW; EDCR670072DCBBA CRC64;

Query Match 89.7%; Score 26; DB 1; Length 710;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
|11111|
DB 379 LEPRSS 384

RESULT 13
CAN3_RAT STANDARD; PRT; 821 AA.
ID CAN3_RAT
AC P16259;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calpain 3 large subunit (EC 3.4.22.17) (Calpain I3) (Calpain p94,
DE large [catalytic] subunit) (Calcium-activated neutral proteinase 3)
DE (CANP 3) (Muscle-specific calcium-activated neutral protease 3 large
DE subunit).
GN CAPN3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=90062125; PubMed=2555341;
RA Sorimachi H., Imajoboh-Ohm S., Emori Y., Kawasaki H., Ohno S.,
RA Minami Y., Suzuki K.;
RT "Molecular cloning of a novel mammalian calcium-dependent protease
RT distinct from both m- and mu-types. Specific expression of the mRNA
RT in skeletal muscle.";
RL J. Biol. Chem. 264:20106-20111(1989).
CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Met-I-Xaa or

```

CC Arg1-Xaa with Leu or Val as the P2 residue.
CC -1- ENZYME REGULATION: Activated by micromolar concentrations of
CC calcium and inhibited by calpastatin.
CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
CC (regulatory) subunit.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: SKELETAL MUSCLE.
CC -1- MISCELLANEOUS: IN RAT THERE SEEMS TO BE 2 TYPES OF CALPAIN:
CC UNIQUTOUS FORMS - CALPAIN I (MICROMOLAR CA++ REQUIRING) AND
CC CALPAIN II (MILI-MOLE CA++ REQUIRING), AND TISSUE SPECIFIC FORMS -
CC CALPAIN P94 AND NC12. THE SMALL UNIT IS COMMON TO ALL FORMS.
CC -1- SIMILARITY: COMPAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2; ALSO KNOWN AS THE
CC CALPAIN FAMILY OF THIOL PROTEASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J05121; AAA41790.1; -.
DR PIR: B34488; B34488.
DR HSSP: P04632; IDKV.
DR MEROPS: C02.004; -.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR001300; Peptidase_C2.
DR InterPro: IPR000169; Thiolprot_act_site.
DR Pfam: PF01067; Calpain_IIT.1.
DR Pfam: PF00036; ehand; 3.
DR Pfam: PF00646; Peptidase_C2; 1.
DR PRINTS: PR00704; CALPAIN.
DR SMART: SM00230; CysPc; 1.
DR SMART: SM00054; Efh; 2.
DR PROSITE: PS00018; EF_HAND; 2.
DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE: PS00639; THIOL_PROTEASE_HIS_FALSE_NEG.
DR PROSITE: PS00640; THIOL_PROTEASE_ASN_FALSE_NEG.
DR Hydrolase: Thiol protease; Calcium-binding; Multigene family; Repeat.
FT DOMAIN 49 234 THIOL PROTEASE DOMAIN I.
FT DOMAIN 235 428 THIOL PROTEASE DOMAIN II.
FT DOMAIN 429 585 DOMAIN III, C2-LIKE DOMAIN.
FT DOMAIN 587 649 LINKER.
FT DOMAIN 650 820 DOMAIN IV.
FT CA_BIND 705 716 EF-HAND 1 (PROBABLE).
FT CA_BIND 735 746 EF-HAND 2 (PROBABLE).
FT DOMAIN 770 811 ANCESTRAL CALCIUM SITE 3 (POTENTIAL).
FT DOMAIN 800 811 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
FT ACT_SITE 129 129 BY SIMILARITY.
FT ACT_SITE 334 334 BY SIMILARITY.
FT ACT_SITE 358 358 BY SIMILARITY.
SQ SEQUENCE 821 AA; 94127 MW; 27FAD2FEA19PBF CRC64;

Query Match 89.7%; Score 26; DB 1; Length 821;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

GN RPOC1.
OS Chlorella vulgaris.
OC Chloroplast.
OC Eukaryota: Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Chlorella.
OX NCBI_TaxID=3077;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=IAM C-27 / TAMIYA;
RX MEDLINE=97303241; PubMed=9159184;
RA Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,
RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,
RA Inamura A., Yoshinaga K., Sugitara M.,
RT "Complete nucleotide sequence of the chloroplast genome from the
RT green alga Chlorella vulgaris: the existence of genes possibly
RT involved in chloroplast division."
RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: IN CHLOROPLAST THE RNA POLYMERASE IS COMPOSED OF FOUR
CC SUBUNITS: ALPHA, BETA, BETA', AND BETA".
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB001684; BAA57970.1; -.
DR InterPro: IPR000722; RNA_POL_A.
DR Pfam: PF00623; RNA_POL_A; 1.
DR transfease; Transcription; DNA-directed RNA polymerase; Chloroplast.
SQ SEQUENCE 836 AA; 96561 MW; FB50D65D2EA7A3DE CRC64;

Query Match 89.7%; Score 26; DB 1; Length 836;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 14
RPOC_CHLVU
ID RPOC_CHLVU STANDARD; PRT; 836 AA.
AC P56300;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6).

```

```

RESULT 15
PEX6_RAT
ID PEX6_RAT STANDARD; PRT; 978 AA.
AC P54777; O55097;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Peroxisome assembly factor-2 (PAF-2) (Peroxisomal-type ATPase 1)
DE (Peroxin-6).
GN PEX6.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=FISCHER 344; TISSUE=Liver;
RX MEDLINE=96083586; PubMed=7493019;
RA Tsukamoto T., Miura S., Nakai T., Yokota S., Shinozawa N.,
RA Suzuki Y., Orl T., Fujiki Y., Sakai F., Bogaki A., Yasuno H.,
RA Osumi T.;
DE "Peroxisome assembly factor-2, a putative ATPase cloned by functional

```

```

RT complementation on a peroxisome-deficient mammalian cell mutant.";
RL Nat. Genet. 11:395-401(1995).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RA Tsukamoto T., Hashiguchi N.;
RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
CC -! FUNCTION: INVOLVED IN PEROXISOME BIOSYNTHESIS. REQUIRED FOR
CC STABILITY OF THE PTS1 RECEPTOR.
CC -! SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN, PEROXISOME.
CC -! SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D63673; BAA09824.1; -.
DR EMBL; D89660; BAA24931.1; -.
DR EMBL; D89657; BAA24931.1; JOINED.
DR EMBL; D89658; BAA24931.1; JOINED.
DR EMBL; D89659; BAA24931.1; JOINED.
DR InterPro; IPR003960; AAA_sub.
DR InterPro; IPR003959; AAA_subfam.
DR Pfam; PF00004; AAA; 2.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00674; AAA; 1.
KW Peroxisome; ATP-binding; Repeat.
FT NP_BIND 740 749 ATP (POTENTIAL).
FT NP_BIND 742 749 ATP (POTENTIAL).
FT MUTAGEN 476 476 K->A: NO LOSS OF FUNCTION.
FT MUTAGEN 748 748 K->A: LOSS OF FUNCTION.
FT CONFLICT 299 299 D -> G (IN REF. 2).
FT CONFLICT 333 333 V -> A (IN REF. 2).
FT CONFLICT 343 343 Q -> R (IN REF. 2).
FT CONFLICT 546 546 R -> C (IN REF. 2).
SO SEQUENCE 978 AA; 104426 MW; F723193B7E95EA97 CRC64;

```

```

Query Match      89.7%; Score 26; DB 1; Length 978;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 LEPRAS 6
    1111:1
Db 935 LEPRSS 940

```

```

Search completed: June 6, 2002, 13:08:09
Job time: 211 sec

```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OK protein - protein search, using sw model

Run on: June 6, 2002, 13:04:13 ; Search time 25.04 Seconds
(without alignments)
41.452 Million cell updates/sec

Title: US-09-627-383-1

Perfect score: 29

Sequence: 1 LEPRAS 6

Scoring table: BLOSUM62

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea: *
2: SP bacteria: *
3: SP fungi: *
4: SP human: *
5: SP invertebrate: *
6: SP mammal: *
7: SP mhc: *
8: SP organelle: *
9: SP phage: *
10: SP plant: *
11: SP rodent: *
12: SP virus: *
13: SP vertebrate: *
14: SP unclassified: *
15: SP virus: *
16: SP bacteriophage: *
17: SP archaea: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	29	100.0	162	16	Q9RLT2
2	29	100.0	242	16	Q9PCP1
3	29	100.0	526	4	Q9NUY2
4	29	100.0	594	10	Q9LGG7
5	29	100.0	635	5	Q24564
6	29	100.0	1060	5	Q46055
7	29	100.0	1077	5	Q9V8H8
8	29	100.0	1397	4	Q9NTG2
9	29	100.0	2359	11	Q9EQ60
10	29	100.0	2548	4	Q9UNJ2
11	29	100.0	2626	11	Q9ZLN3
12	27	93.1	311	2	Q68976
13	27	93.1	356	16	Q986C9
14	27	93.1	357	7	Q9S560
15	27	93.1	357	7	Q9S410
16	27	93.1	357	7	Q9S411

17	27	93.1	366	16	Q92YG1	Q92YG1 rhizobium m
18	27	93.1	385	16	Q92KK9	Q92KK9 rhizobium m
19	27	93.1	422	16	Q9KSV9	Q9KSV9 vibrio chol
20	27	93.1	887	16	Q98EH3	Q98EH3 rhizobium l
21	26	89.7	104	4	Q9POE3	Q9POE3 homo sapien
22	26	89.7	147	11	Q925S3	Q925S3 mus musculu
23	26	89.7	153	2	Q44626	Q44626 bruceella ab
24	26	89.7	170	11	Q925S2	Q925S2 mus musculu
25	26	89.7	174	16	Q989Y3	Q989Y3 rhizobium l
26	26	89.7	177	16	Q92T87	Q92T87 rhizobium m
27	26	89.7	214	16	Q9RYS3	Q9RYS3 deinooccus
28	26	89.7	252	3	Q9HF62	Q9HF62 ashbya goss
29	26	89.7	279	16	Q984S7	Q984S7 rhizobium l
30	26	89.7	321	16	Q9RZ7	Q9RZ7 deinooccus
31	26	89.7	335	4	Q96T06	Q96T06 homo sapien
32	26	89.7	339	10	Q9XES7	Q9XES7 glycine max
33	26	89.7	348	10	Q9C9M7	Q9C9M7 arabidopsis
34	26	89.7	352	16	Q94GZ7	Q94GZ7 oryza sativ
35	26	89.7	352	16	Q9I763	Q9I763 pseudomonas
36	26	89.7	360	2	Q9FA57	Q9FA57 azarcus ev
37	26	89.7	406	2	Q93CM9	Q93CM9 helicobacte
38	26	89.7	407	4	Q96LW0	Q96LW0 homo sapien
39	26	89.7	477	3	Q42968	Q42968 schizosach
40	26	89.7	479	4	Q9NZI6	Q9NZI6 homo sapien
41	26	89.7	481	11	Q99PF2	Q99PF2 mus musculu
42	26	89.7	493	11	Q9RLI3	Q9RLI3 rattus norv
43	26	89.7	547	5	Q9N637	Q9N637 leishmania
44	26	89.7	562	10	Q9S7I3	Q9S7I3 arabidopsis
45	26	89.7	601	5	Q17402	Q17402 caenorhabd

ALIGNMENTS

RESULT 1
ID Q9RLT2 PRELIMINARY: PRT: 162 AA.
AC Q9RLT2;
DT 01-MAY-2000 (TRMBLrel. 13, Created)
DT 01-MAY-2000 (TRMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TRMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 17.0 KDA PROTEIN.
GN DR1748.
OS Deinooccus radiodurans.
OC Bacteria; Thermus/Deinooccus group; Deinococcales; Deinooccus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RL;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D., Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L., Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M., Vamathevan J.J., Lam P., McDonald L., Uitterback T., Zalewski C., Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D., Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C., Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinooccus radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL: AE002016; AAF11308.1; -
DR TIGR: DR1748; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 162 AA; 16986 MW; 3023A45A174A7BF4 CRC64;

Query Match 100.0%; Score 29; DB 16; Length 162;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LEPRAS 6
Db 151 LEPRAS 156

```

RESULT 2
Q9PCP1 PRELIMINARY; PRT; 242 AA.
AC Q9PCP1;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN XF1737.
GN XF1737.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordn S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Fromme M., Furian L.R.,
RA Gardiner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.V., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,
RA Quaglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,
RA de Souza A.P., Terezzi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL: AEO03997; AAF84546.1; -
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 242 AA; 25385 MW; 857F70152F2B0C1B CRC64;

```

Query Match 100.0%; Score 29; DB 16; Length 242;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 LEPRAS 6
Db 75 LEPRAS 80

```

```

RESULT 3
Q9NUY2 PRELIMINARY; PRT; 526 AA.
AC Q9NUY2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CDNA FL11061 FIS, CLONE PLACE1004777, WEAKLY SIMILAR TO
DE N-CHIMAERIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Isoqal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahara K., Masuno Y., Sasaki N.;
RT "NEDO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK001923; BAA91979.1; -
DR HSSP: 007960; IRGP.
DR InterPro: IPR000198; RHO GAP.
DR SMART: SM00324; RHO GAP; 1.
SQ SEQUENCE 526 AA; 59212 MW; 868624871E4C860C CRC64;

```

Query Match 100.0%; Score 29; DB 4; Length 526;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 LEPRAS 6
Db 335 LEPRAS 340

```

```

RESULT 4
Q9LGZ7 PRELIMINARY; PRT; 594 AA.
AC Q9LGZ7;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SIMILAR TO ANABIDOPSIS THALIANA CHROMOSOME 4 BAC T15B16.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(CA3) genomic DNA, chromosome 1, PAC
RL clone: p0469E05."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002480; BAA96574.1; -
DR InterPro: IPR003657; WRKY.
SQ SEQUENCE 594 AA; 61071 MW; 7A1D4E5640448789 CRC64;

```

Query Match 100.0%; Score 29; DB 10; Length 594;
 Best Local Similarity 100.0%; Pred. No. 11e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 LEPRAS 6
Db 218 LEPRAS 223

```

```

RESULT 5
Q24564 PRELIMINARY; PRT; 635 AA.
AC Q24564;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE MER PROTEIN (CYTOSKELETAL PROTEIN).
GN MER OR EWR2 OR DMERLIN OR CG14228.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.

```

OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Aydayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris J.M., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA de Pablo B., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jastli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mantei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mervulov G., Milshina N.V., Modarri C., Morris J., Mostrel A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Relbert K., Remington K., Sanders R.D.C., Scheeler F., Shen H.,
 RA Shu B.C., Siden-Kimios I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodgett T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.).
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96234672; PubMed=8666669;
 RA McCartney B.M., Fehon R.G.;
 RT "Distinct cellular and subcellular patterns of expression imply
 RT distinct functions for the *Drosophila* homologues of moesin and the
 RT neurofibromatosis 2 tumor suppressor, merlin."
 RL J. Cell Biol. 133:843-852(1996).
 RN [3]
 RP SEQUENCE OF 218-307 FROM N.A.
 RA Winge P., Fleming J.T., Gobel V.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE003512; AAF49005.1; -
 DR EMBL: U49724; AAB08449.1; -
 DR EMBL: U23799; AAB65060.1; -
 DR FlyBase: FBgn0013951; Mer.
 DR InterPro: IPR000299; Band_4.1.
 DR InterPro: IPR000798; Ezrin_radixin_moesin.
 DR Pfam: PF00373; Band_41; 1.
 DR Pfam: PF00769; ERN; 1.
 DR PRINTS: PR00935; BAND41.
 DR SMART: SM00295; B41; 1.
 DR PROSITE: PS50057; BAND_41_3; 1.
 DR PROSITE: PS50057; BAND_41_3; 1.
 SQ SEQUENCE 635 AA; 74492 MW; 940D6B8A0D160A3F CRC64;

Query Match 100.0%; Score 29; DB 5; Length 635;
 Best Local Similarity 100.0%; Pred. NO. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
 DB 27 LEPRAS 32
 ID 046055 PRELIMINARY; PRT; 1060 AA.
 AC 046055;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE EG:165H7.3 PROTEIN.
 GN L(1)IBB OR EG:165H7.3 OR CG3923.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA de Pablo B., Madueno E., Modolell J.;
 RT "Sequencing the distal X chromosome of *Drosophila melanogaster*."
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Benos P.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL009188; CA15674.2; -
 DR FlyBase: FBgn0001337; 1(1)IBb.
 DR InterPro: IPR001494; IBB_NT.
 SQ SEQUENCE 1060 AA; 119276 MW; 650B030D25B9156 CRC64;

Query Match 100.0%; Score 29; DB 5; Length 1060;
 Best Local Similarity 100.0%; Pred. NO. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
 DB 5 LEPRAS 10
 ID 09V8H8 PRELIMINARY; PRT; 1077 AA.
 AC 09V8H8;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE CG12263 PROTEIN.
 GN CG12263.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Aydayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris J.M., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Palos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Doudson K., Dup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dushin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Idegawa C.,
RA Jatali M., Kailush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Mishina N.V., Moharir C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacled J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Sampson M., Skupski M.P., Smith T.,
RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gbbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003799; AAF57689.1;
DR FlyBase: FBgn0034346; CG12263.
SQ SEQUENCE 1077 AA; 121329 MW; 9EB94BF2E243F5 CRC64;

Query Match 100.0%; Score 29; DB 5; Length 1077;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
|||||
Db 828 LEPRAS 833

RESULT 8
09NTG2 PRELIMINARY; PRT; 1397 AA.
AC 09NTG2;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE HYPOHETICAL 157.9 KDA PROTEIN (FRAGMENT).
GN DKFZP434F1016.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TESTIS;
RA Koehler K., Beyer A., Mewes H.W., Gassenhuber J., Wismann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL137287; CAB0679.1;
DR HSSP: Q07960; IRGP.
DR InterPro: IPR002219; DAG_PE-Bind.
DR InterPro: IPR000198; RhogAP.
DR SMART: SM00109; C1; 1.
DR SMART: SM00324; RhogAP; 1.
DR PROSITE: PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR PROSITE: PS50081; DAG_PE_BIND_DOM_2; 1.
KM Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 1397 AA; 157884 MW; 0FBB877C358526 CRC64;

Query Match 100.0%; Score 29; DB 4; Length 1397;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
|||||
Db 1206 LEPRAS 1211

RESULT 9
09EQ60 PRELIMINARY; PRT; 2359 AA.
AC 09EQ60;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE CALCIUM CHANNEL ALPHA-1-H SUBUNIT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=21264893; PubMed=11073957;
RA McEry J.E., Santi C.M., Hamming K.S.C., Mezeyova J., Sutton K.G.,
RA Baillie D.L., Stea A.A., Snutch T.P.;
RT "Molecular and functional characterization of a family of rat brain T-
type calcium channels.";
RL J. Biol. Chem. 276:3999-4011(2001).
DR EMBL: AF290213; AAC35187.1;
DR InterPro: IPR000636; Calton_chan_non_11g.
DR InterPro: IPR002111; Cal_channel_Tyrl.
DR InterPro: IPR002077; Ca_channel.
DR InterPro: IPR001682; Channel_pore_Ca_Na.
DR InterPro: IPR003915; PKD_2.
DR Pfam: PF00520; Ion_trans_4.
DR PRINTS: PR00167; CACHANNEL.
DR PRINTS: PR01433; POLYCYSTIN2.
SQ SEQUENCE 2359 AA; 261138 MW; F738083E94180081 CRC64;

Query Match 100.0%; Score 29; DB 11; Length 2359;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
|||||
Db 1969 LEPRAS 1974

RESULT 10
09UNJ2 PRELIMINARY; PRT; 2548 AA.
AC 09UNJ2;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE MYOSIN-IXA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99339979; PubMed=10409426;
RA Gorman S.W., Halder N.B., Grieshammer U., Swiderski R.E., Kim E.,
RA Welch J.W., Seaby C., Ieng S., Carni R., Sheffield V.C., Duhl D.M.;
RT "The cloning and developmental expression of unconventional myosin IXA
(MYO9A) a gene in the Bardet-Biedl syndrome (BBS4) region at
chromosome 15q22-q23.";
RL Genomics 59:150-160(1999).
DR EMBL: AF117888; AAD49195.1;
DR HSSP: P08799; IMND.
DR InterPro: IPR002219; DAG_PE-Bind.

DR InterPro: IPR000048; IO.
 DR InterPro: IPR001609; myosin_head.
 DR InterPro: IPR000159; RA.
 DR InterPro: IPR000198; RhogAP.
 DR Pfam: PF00130; DAG_Pe-bind; 1.
 DR Pfam: PF00612; IO; 5.
 DR Pfam: PF00063; myosin_head; 2.
 DR Pfam: PF00788; RA; 1.
 DR Pfam: PF00620; RhogAP; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 2.
 DR SMART: SM00109; C1; 1.
 DR SMART: SM00015; IO; 5.
 DR SMART: SM00242; MYSC; 1.
 DR SMART: SM00314; RA; 1.
 DR SMART: SM00324; RhogAP; 1.
 DR PROSITE: PS00479; DAG_Pe_BIND_DOM_1; UNKNOWN_1.
 DR PROSITE: PS50081; DAG_Pe_BIND_DOM_2; 1.
 DR PROSITE: PS50096; IO; 1.
 SO SEQUENCE 2548 AA; 292703 MW; B93E76C2A0E9A356 CRC64;

Query Match 100.0%; Score 29; DB 4; Length 2548;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
 DB 2357 LEPRAS 2362

RESULT 11

ID 0921N3 PRELIMINARY; PRT; 2626 AA.

AC 0921N3; 01-MAY-1999 (TREMBlrel. 10, Created)

DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)

DE 01-DEC-2001 (TREMBlrel. 19, last annotation update)

DE MYOSIN-RHOGAP PROTEIN, MYR 7.

GN MYOA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.
 RA Chieragati E., Gaertner A., Stoeffler H.E., Baehler M.,
 RT "Myr 7 is a novel myosin-Rho-GAP molecule expressed in rat brain."

RT Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.

DR EMBL: AJ001713; CAA04946.1; -

DR HSSP: P08799; 1MND.

DR InterPro: IPR002106; AA_trna_ligase-II.

DR InterPro: IPR002219; DAG_Pe-bind.

DR InterPro: IPR000048; IO.

DR InterPro: IPR001609; myosin_head.

DR InterPro: IPR000159; RA.

DR InterPro: IPR00198; RhogAP.

DR Pfam: PF00130; DAG_Pe-bind; 1.

DR Pfam: PF00612; IO; 5.

DR Pfam: PF00063; myosin_head; 2.

DR Pfam: PF00788; RA; 1.

DR Pfam: PF00620; RhogAP; 1.

DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 2.
 DR SMART: SM00109; C1; 1.
 DR SMART: SM00015; IO; 5.
 DR SMART: SM00242; MYSC; 1.
 DR SMART: SM00314; RA; 1.
 DR SMART: SM00324; RhogAP; 1.
 DR PROSITE: PS00479; AA_trna_LIGASE-II_1; UNKNOWN_1.
 DR PROSITE: PS00479; DAG_Pe_BIND_DOM_1; UNKNOWN_1.
 DR PROSITE: PS50081; DAG_Pe_BIND_DOM_2; 2.
 SO SEQUENCE 2626 AA; 301378 MW; 3F70610271E4D791 CRC64;

Query Match 100.0%; Score 29; DB 11; Length 2626;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
 DB 2443 LEPRAS 2448

RESULT 12

ID 068976 PRELIMINARY; PRT; 311 AA.

AC 068976; 01-AUG-1998 (TREMBlrel. 07, Created)

DT 01-AUG-1998 (TREMBlrel. 07, last sequence update)

DE 01-DEC-2001 (TREMBlrel. 19, last annotation update)

DE HYPOTHETICAL 35.1 KDA PROTEIN (FRAGMENT).

OS Brucella melitensis.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Brucellaceae; Brucella.

OX NCBI_TaxID=29459;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-133;

RA Hernandez-Castro R., Sahagun-Ruiz A., Verdugo-Rodriguez A.,
 RA Waghela S., Gutierrez-Pabellio A.J., Adams A.G., Suarez-Guemes F.,
 RT "Cloning and sequence of hypothetical protein of Brucella
 melitensis."

RL Submitted (APR-1998) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF055568; AAC14571.1; -

DR InterPro: IPR002787; DUF85.

DR Pfam: PF01932; DUF85; 1.

DR ProDom: PD011382; DUF85; 1.

KT Hypothetical protein.

FM NON_TER

SO SEQUENCE 311 AA; 35078 MW; B523B28790C7DA6B CRC64;

Query Match 93.1%; Score 27; DB 2; Length 311;
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEPRAS 6
 DB 173 MEPRAS 178

RESULT 13

ID 0986C9 PRELIMINARY; PRT; 356 AA.

AC 0986C9; 01-OCT-2001 (TREMBlrel. 18, Created)

DT 01-OCT-2001 (TREMBlrel. 18, last sequence update)

DE 01-OCT-2001 (TREMBlrel. 18, last annotation update)

DE MLT7A17 PROTEIN.

GN MLT7A17.

OS Rhizobium loti (Mesorhizobium loti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Phyllobacteriaceae; Mesorhizobium.

OX NCBI_TaxID=381;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MAFE303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Tachizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.,
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti."

RL DNA Res. 7:331-338(2000).
DR EMBL; AP003011; BAB53524.1; -.
DR InterPro: IPR002787; DUF85.
DR Pfam; PF01932; DUF85; 1.
DR ProDom: PD011382; DUF85; 1.
KW Complete proteome.
SQ SEQUENCE 356 AA; 40498 MW; F7437E1708E175A4 CRC64;

Query Match 93.1%; Score 27; DB 16; Length 356;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
:|||||
Db 66 MEPRAS 71

RESULT 14
ID Q95560 PRELIMINARY; PRT; 357 AA.
AC Q95560;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MAJOR HISTOCOMPATIBILITY COMPLEX CLASS I ANTIGEN.
OS Peromyscus maniculatus (Deer mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
OC Peromyscus.
OX NCBI_Taxid=10042;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
FX MEDLINE=96269997; Pubmed=8662082;
RA Crew M.D., Bates L.M., Douglass C.A., York J.L.;
RT "Expressed Peromyscus maniculatus (Pema) MHC class I genes:
RT evolutionary implications and the identification of a gene encoding a
RT Qa1-like antigen.";
RL Immunogenetics 44:177-185(1996).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICRoglobulin) (BY SIMILARITY).
CC EMBL; U12887; AAB17696.1; -.
DR HSSP; P01900; IBI1.
DR InterPro: IPR003597; Iq-cl.
DR InterPro: IPR003006; Iq_MHC.
DR InterPro: IPR001039; MHC_I.
DR Pfam; PF00047; Iq; 1.
DR Pfam; PF00129; MHC_I; 1.
DR ProDom: PD000050; MHC_I; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Glycoprotein; Transmembrane.
SQ SEQUENCE 357 AA; 40223 MW; 0618C51F04B31560 CRC64;

Query Match 93.1%; Score 27; DB 7; Length 357;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
:|||||
Db 66 MEPRAS 71

RESULT 15
ID Q95410 PRELIMINARY; PRT; 357 AA.
AC Q95410;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MHC CLASS I (FRAGMENT).
OS Hylobates lar (Common gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylodactylidae; Hylobates.
OX NCBI_Taxid=9580;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92218857; Pubmed=1560209;
RA Chen Z.W., McAdam S.N., Hughes A.L., Dogon A.L., Letvin N.L.,
RA Watkins D.I.;
RT "Molecular cloning of orangutan and gibbon MHC class I CDNA. The HLA-A
RT and -B loci diverged over 30 million years ago.";
RL J. Immunol. 148:2547-2554(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Chen Z.W., McAdam S.N., Hughes A.L., Watkins D.I.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICRoglobulin) (BY SIMILARITY).
CC EMBL; U50089; AAB08072.1; -.
DR HSSP; O19673; IHSB.
DR InterPro: IPR003597; Iq-cl.
DR InterPro: IPR003006; Iq_MHC.
DR InterPro: IPR001039; MHC_I.
DR Pfam; PF00047; Iq; 1.
DR Pfam; PF00129; MHC_I; 1.
DR ProDom: PD000050; MHC_I; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Glycoprotein; Transmembrane.
FT NON_TER 1
SQ SEQUENCE 357 AA; 40026 MW; 682DDFBEB8DB9F361 CRC64;

Query Match 93.1%; Score 27; DB 7; Length 357;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LEPRAS 6
:|||||
Db 61 MEPRAS 66

Search completed: June 6, 2002, 13:07:53
Job time: 220 sec

THIS PAGE BLANK (USPTO)